

PA 1245139

REC'D 01 FEB 2005

WIPO

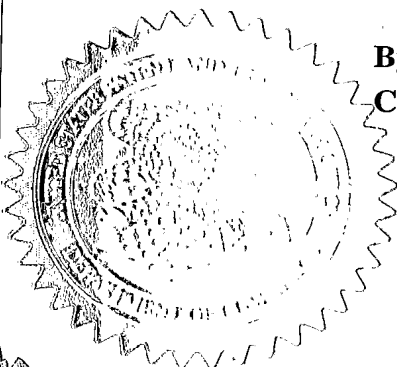
PCT

THE UNITED STATES OF AMERICA**TO ALL TO WHOM THESE PRESENTS SHALL COME:****UNITED STATES DEPARTMENT OF COMMERCE****United States Patent and Trademark Office****November 09, 2004**

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 60/538,799**FILING DATE: January 23, 2004****PRIORITY
DOCUMENT****SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)**

**By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS**



M. K. Hawkins
M. K. HAWKINS
Certifying Officer

16669 U.S. PTO
012304

SUBSTITUTE for Provisional Application for Patent Cover Sheet PTO/SB/16 (08-03)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

22154 U.S. PTO
60/538799
012304

DOCKET NUMBER		ITR0048PV	
INVENTOR(S)			
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)	
Stefano Alfredo Elisabetta Agostino Bruno Annalisa	Colloca Nicosia Sporeno Cirillo Bruni Ercole Meola	Rome, Italy 00153 Rome, Italy 00144 Rome, Italy 00153 Lanuvio, Italy 00040 Pomezia, Italy 00040 Ariccia, Italy 00040	
<input type="checkbox"/> Additional inventors are being named on the separately numbered sheets attached hereto			
TITLE OF THE INVENTION (500 characters max)			
CHIMPANZEE ADENOVIRUS VACCINE CARRIERS			
CORRESPONDENCE ADDRESS			
Direct all Correspondence to: Merck & Co., Inc. Patent Department - RY60-30 P.O. Box 2000 Rahway			
STATE	New Jersey	ZIP CODE	07065
		COUNTRY	U.S.A.
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification	Number of Pages	64	<input type="checkbox"/> CD(s), Number
<input checked="" type="checkbox"/> Drawing(s)	Number of Sheets	153	<input type="checkbox"/> Other (specify)
<input type="checkbox"/> Application Data Sheet. See 37 CFR 1.76			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)			
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees		FILING FEE AMOUNT (\$)	\$160.00
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 13-2755			

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE Patricia Chisholm
 TYPED or PRINTED NAME Patricia Chisholm
 TELEPHONE 732-594-5738

Date 01/23/2004

REGISTRATION NO. 45,822
(if appropriate)

NOTE: Mail to Mail Stop Provisional Application

EXPRESS MAIL CERTIFICATE	
DATE OF DEPOSIT	January 23, 2004
EXPRESS MAIL NO.	EV 32315351505
I HEREBY CERTIFY THAT THIS CORRESPONDENCE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE AS EXPRESS MAIL "POST OFFICE TO ADDRESSEE" ON THE ABOVE DATE IN AN ENVELOPE ADDRESSED TO COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450.	
MAILED BY	DATE
Here Schepers	Jan 23, 2004

In Duplicate

CHIMPANZEE ADENOVIRAL VACCINE CARRIERS

FIELD OF THE INVENTION

5 The present invention relates to the field of recombinant vectors and more specifically to the production and use of recombinant replication-defective chimpanzee adenoviral vectors to elicit immune responses in mammalian hosts.

BACKGROUND OF THE INVENTION

10 The adenoviruses (Ads) comprise a large family of double-stranded DNA viruses found in amphibians, avians, and mammals which have genome organization and a nonenveloped icosahedral capsid structure (Straus, Adenovirus infections in humans. In *The Adenoviruses*. 451-498, 1984; Hierholzer *et al.*, *J. Infect. Dis.*, 158: 804-813, 1988; Schnurr and Dondero, *Intervirology*, 36: 79-83, 1993; Jong *et al.*, *J Clin Microbiol.*, 37:3940-3945:1999). In
15 contrast to retroviruses, adenoviruses can transduce numerous cell types of several mammalian species, including both dividing and nondividing cells, without integrating into the genome of the host cell.

 Generally speaking, adenoviral DNA is typically very stable and remains episomal (e.g., extrachromosomal), unless transformation or tumorigenesis has occurred. In addition,
20 adenoviral vectors can be propagated to high yields in well-defined production systems which are readily amenable to pharmaceutical scale production of clinical grade compositions. These characteristics and their well-characterized molecular genetics make recombinant adenoviral vectors good candidates for use as vaccine carriers. Typically, the production of recombinant adenoviral vectors relies on the use of a packaging cell line which is capable of complementing
25 the functions of adenoviral gene products that have been either deleted or engineered to be nonfunctional.

 Presently, two well-characterized human subgroup C adenovirus serotypes (i.e., hAd2 and hAd5) are widely used as the sources of the viral backbone for most of the adenoviral vectors that are used for gene therapy. Replication-defective human adenoviral vectors have also
30 been tested as vaccine carriers for the delivery of a variety of immunogens derived from a variety of infectious agents (e.g., viruses, parasites, or bacterial pathogens) and tumor cells, including tumor-associated antigens (TAAs). Studies conducted in experimental animals (e.g., rodents, canines and nonhuman primates) indicate that recombinant replication-defective human adenoviral vectors carrying transgenes encoding immunogens derived from the E6 and E7
35 oncoproteins of human papillomavirus (HPV-16) (He, Z *et al.*, (2001) *Virology*, 270:3583-3590,

the rabies virus glycoprotein (Xiang, Z. *et al* (1996) *Virology*, 219:220-227), the circumsporozoite protein of *Plasmodium falciparum* Rodriguez, E. *et al.* (1997) *J. Immunol.* 158:1268-1274) as well as other heterologous antigens elicit both humoral and cell-mediated immune responses against the transgene product. Generally speaking, investigators have reported success using human adenoviral vectors as vaccine carriers in nonhuman experimental systems by either using an immunization protocols that utilizes high doses of recombinant adenoviral vectors that are predicted to elicit immune responses; or by using immunization protocols which employ the sequential administration of adenoviral vectors that are derived from different serotypes but which carry the same transgene product as boosting immunizations (Mastrangeli, *et al.*, *Human Gene Therapy*, 7: 79-87 (1996).

However, it is predicted that vaccine carriers derived from ubiquitous human serotypes, such as types 2 and 5, will encounter preexisting humoral and cellular immunity in the human population. Thus, although replication-defective recombinant human adenoviruses have been successfully employed as vaccine carriers in experimental systems employing rodent, canine, and nonhuman primate hosts; human innate and adaptive immunity is expected to significantly limit the utility of these serotypes as vaccine carriers. This expectation is based on the fact that subgroup C, which includes type 2 and type 5, adenoviral infection is endemic in the human population. As a consequence, the majority of humans seroconvert within the first five years of life as the result of a natural infection. Thus, vectors derived from viruses that naturally infect and replicate in humans may not be optimal candidates for use as vaccine carriers.

Another problem associated with the use of human adenoviral-derived vectors is the risk that the production method used to propagate the recombinant viruses will give rise to vector stocks that are contaminated with replication competent adenovirus (RCA). This is caused by homologous recombination between overlapping sequences from the recombinant vector and the adenoviral genes that are present in the E1-complementing helper cell lines such as human 293 (Graham, F.L. *et al.*, (1977) *J. Gen. Virol.* 36:59-72.) cells. The presence of RCA in vector stocks prepared for use in clinical trials constitutes a safety risk because it can promote the mobilization and spread of the replication defective virus. Spread of the defective virus can aggravate the host immune response and cause other adverse immunopathological consequences (Fallux, F. J., *et al.* *Human Gene Therapy* 9: 1909-1917 (1998). Accordingly, the Food and Drug Administration (FDA) and other regulatory bodies have promulgated guidelines which establish limits on the levels of RCA that can be present in vector preparations intended for clinical use. The intent of imposing RCA limits is to ensure limited exposure of patients to replicating adenovirus in compositions that are used in clinical trials.

Thus, there continues to be a need for the development of adenoviral vaccine carriers that are suitable for use in mammalian hosts which are: easy to manipulate, amenable to pharmaceutical scale production and long term storage, capable of high-level replication in human complementation cell lines, highly immunogenic, devoid of neutralizing B cell epitopes that cross-react with the common serotypes of human adenoviruses, comply with the safety RCA standards promulgated by regulatory agencies, and which are amenable for use in prime/boost protocols that are suitable for use in humans.

SUMMARY OF THE INVENTION

The present invention relates to recombinant replication-defective adenovirus vectors derived from chimpanzee adenoviruses and methods for generating chimpanzee adenoviral vectors in human E1-expressing cell lines. The invention also provides methods for generating clinical grade vector stocks suitable for use in humans and means for using the disclosed vectors as vaccine carriers to elicit protective and/or therapeutic immune responses. The invention further provides methods for using the recombinant adenoviruses of the invention to prepare vaccine compositions designed to delivery, and direct the expression of, transgenes encoding immunogens. In one embodiment, the invention contemplates the use of the disclosed vectors as vaccine carriers for the administration of vaccines comprising transgenes encoding immunogens derived from an infectious agent. In a second embodiment, the invention contemplates the use of the disclosed vectors to prepare and administer cancer vaccines. In a particular embodiment, the invention contemplates the preparation and administration of a cancer vaccine comprising a transgene encoding a TAA.

In one aspect, the invention discloses the complete genomic sequence of five chimpanzee adenoviruses (ChAds), referred to herein as ChAd3 (SEQ ID NO: 1) (Figures 5A-5V), ChAd6 (SEQ ID NO: 2) (Figures 6A-6V, CV32 (SEQ ID NO:3) (Figures 7A-7K), CV33 (SEQ ID NO: 4) (Figures 8A-8K), and CV23 (SEQ ID NO:5) (Figures 9A-9J). ChAd3 and ChAd6 represent novel adenoviruses isolated according to the methods disclosed herein. The genomes of the ChAd3 and ChAd6 are 37741 and 36648 base pairs in length, respectively. ChAd3 hexon gene is comprised between nt 19086-21968 (SEQ ID NO: 41) while fiber gene is comprised between nt 32805-34490 (SEQ ID NO: 42). ChAd6 hexon gene is comprised between nt 18266-21127 (SEQ ID NO: 43) while fiber gene is comprised between nt 32218-33555 (SEQ ID NO: 44). Based on sequence homology deduced from a multiple sequence alignment of full-length hexon peptides, ChAd3 has been classified into human subgroup C and ChAd6 has been classified into human subgroup E.

The genomes of the CV32, CV33 and CV23 adenoviruses are 36,606, 36,535, and 32,020 base pairs in length, respectively. CV32 (Pan 6, ATCC N. VR-592), CV33 (Pan 7, ATCC N. VR-593) and CV23 (Pan 5) (Esoterix Inc.) have all been determined to be related to human Ad4 (hAd4) (subgroup E) (Wigand, R *et al. Intervirology* 1989, 30:1-9). However, based on hexon sequence alignment CV32 has subsequently characterized as being more closely analogous to human subgroup D members than to hAd4.

In a second aspect, the invention provides nucleotide sequences for the fiber and hexon genes of ten additional chimpanzee adenoviruses (ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17 and ChAd19) isolated according to the methods disclosed herein. The fiber gene sequences are set forth in Figures 10-19 and SEQ ID NOS: 6-15: (SEQ ID NO: 6, ChAd20); (SEQ ID NO: 7, ChAd4); (SEQ ID NO: 8, ChAd5); (SEQ ID NO: 9, ChAd7); (SEQ ID NO: 10, ChAd9); (SEQ ID NO: 11, ChAd10); (SEQ ID NO: 12, ChAd11); (SEQ ID NO: 13, ChAd16) (SEQ ID NO: 14, ChAd17) and (SEQ ID NO: 15, ChAd19). Figures 20A-20D provides a comparison of the amino acid sequences of the fiber proteins disclosed and claimed herein.

The hexon gene sequences are set forth in Figures 21-30 and SEQ ID NOS: 16-25: (SEQ ID NO: 16, ChAd20); (SEQ ID NO: 17, ChAd4); (SEQ ID NO: 18, ChAd5); (SEQ ID NO: 19, ChAd7); (SEQ ID NO: 20, ChAd9); (SEQ ID NO: 21, ChAd10); (SEQ ID NO: 22, ChAd11); (SEQ ID NO: 23, ChAd16); (SEQ ID NO: 24, ChAd17) and (SEQ ID NO: 25, ChAd19). Figures 31A-31M provide a comparison of the amino acid sequences of the hexon proteins disclosed and claimed herein. A multiple sequence alignment of hexon proteins allows an artisan to perform a phylogenetic analysis of that is consistent with the proposed classification of human adenoviral serotypes (Rux, J.J., *et al* (2003) *J. Virol.* 77:9553-9566).

In an alternative aspect, the invention further provides ten additional chimpanzee adenovirus isolates. Samples comprising ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17 and ChAd19 were deposited on December 12, 2003 with the European Collection of Cell Cultures (ECACC, Porton Down, Salisbury, Wiltshire, SP4 0JG, United Kingdom) as an original deposit under the Budapest Treaty. The deposits were assigned accession numbers: 03121201 (ChAd4), 03121202 (ChAd5), 03121203 (ChAd7), 03121204 (ChAd9), 03121205 (ChAd10), 03121206 (ChAd11), 03121207 (ChAd16), 03121208 (ChAd17), 03121209 (ChAd19) and 03121210 (ChAd20). These deposits will be maintained under the terms of the *Budapest Treaty* on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. These deposits were made merely as a convenience for those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. All restrictions on the availability to the public of the deposited material will be

irrevocably removed, except for the requirements specified in 37 C.F.R. §1.808(b), upon the granting of a patent.

In a third aspect, the invention provides replication-defective recombinant adenoviral vectors which are capable of infecting mammalian cells, preferably human cells, and directing expression of encoded transgene product(s). As demonstrated herein, the disclosed vectors are suitable for use as vaccine carriers for the delivery of transgenes comprising immunogens against which an immune response is desired. In particular embodiments, the invention provides recombinant replication-defective chimpanzee adenoviral vectors that are capable of high-level replication in human E1-expressing (i.e., packaging) cell lines. In one embodiment, the invention provides recombinant adenoviruses that are capable of replicating in PER.C6™ cells.

Generally speaking, the recombinant vectors encompassed by the invention provide vaccine carriers that will evade pre-existing immunity to the adenovirus serotypes that are typically encountered in the human population. More specifically, the recombinant vectors of the invention comprise vector backbone sequences which are shown herein to be devoid of neutralizing B epitopes that cross-react with the common serotypes of human adenoviral derived vectors.

The invention further provides group-specific shuttle vectors that include an adenoviral portion and a plasmid portion, wherein said adenoviral portion generally comprises: a) viral left end (ITR and packaging signal), part of the pIX gene and viral genome right end; and b) a gene expression cassette. The group-specific shuttle vectors are designed to exploit the nucleotide sequence homology which is observed between adenoviruses that are assigned to the same serotype subgroup (i.e., subgroups A, B, C, D or E), and can be used to manipulate the nucleotide sequences disclosed herein and/or to clone other chimpanzee adenoviruses belonging to the same subgroup generating an adenovirus pre-plasmid containing a chimp adenoviral genome deleted of E1 region.

Other aspects of this invention include host cells comprising the adenoviral vaccine vectors and/or the adenovirus pre-plasmid vectors, methods of producing the vectors comprising introducing the adenoviral vaccine vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant adenoviral vaccine vectors. In a particular embodiment, the invention provides a method of producing a replication-defective chimpanzee adenoviral vector comprising introducing one of the disclosed adenoviral vectors into an adenoviral E-1 expressing human cell, and harvesting the resulting recombinant adenoviruses.

Another aspect of the invention also provides vaccine compositions which comprise an adenoviral vector of the invention. Compositions comprising recombinant

chimpanzee adenoviral vectors may be administered alone or in combination with other viral- or non-viral-based DNA/protein vaccines. They also may be administered as part of a broader treatment regimen. These compositions can be administered to mammalian hosts, preferably human hosts, in either a prophylactic or therapeutic setting. As shown herein, administration of the disclosed vaccine compositions, either alone or in a combined modality, such as a prime boost regimen or multiple injections of serologically distinct Ad vectors results in the induction of an immune response in a mammal that is capable of specifically recognizing the immunogen encoded by the transgene.

One of the methods disclosed and claimed herein, comprises administering to a mammal (that is either naïve or primed to be immunoreactive to a target antigen), a sufficient amount of a recombinant chimpanzee adenoviral vector, containing at least a functional deletion of its wild-type E1 gene, carrying a sequence comprising a promoter capable of directing expression of a nucleotide sequence encoding the least one target antigen, wherein administration of the recombinant vector elicits (or primes) an antigen-specific immune response.

In one embodiment, the invention provide a method designed to induce an immune response (prophylactic or therapeutic) against an infectious agent (e.g., a viral or bacterial pathogen or a mammalian parasite). In a second embodiment, the invention provides a method designed to induce an immune response in a mammal that will break tolerance to a self-antigen, such as a TAA. This aspect of the invention contemplates the use of the disclosed vectors as a vaccine carrier for the preparation and administration of cancer vaccines.

Yet other embodiments and advantages of the present invention will be readily apparent from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic drawing which summarizes the cloning strategy used to construct a ChAd6 shuttle vector (pARS ChAd6-3).

Figure 2 is a schematic drawing which illustrates the cloning strategy used to clone the ChAd6 viral genome by homologous recombination in *E.coli* strain BJ5183.

Figure 3 is a schematic drawing illustrating the elements of various ChAd6 shuttle plasmids including: pARS ChAd6-3 GAG; pARS ChAd6-3 SEAP; pARS ChAd6-3 EGFP; and pARS ChAd6-3 NS MUT.

Figure 4 is a schematic drawing which illustrates the homologous recombination scheme utilized to clone the ChAd6 $\Delta E1$ expression vectors.

Figures 5A-5V provides the genomic nucleotide sequence of ChAd3 (SEQ ID NO: 1).

Figures 6A-6V provides the genomic nucleotide sequence of ChAd6 (SEQ ID NO: 2).

Figures 7A-7K provides the genomic nucleotide sequence of CV32 (SEQ ID NO: 3).

Figures 8A-8K provides the genomic nucleotide sequence of CV33 (SEQ ID NO: 4).

Figures 9A-9J provides the genomic nucleotide sequence of CV23 (SEQ ID NO: 5).

Figures 10 A and B provides the nucleotide sequence of the fiber gene of ChAd20 (SEQ ID NO: 6).

Figures 11 A and B provides the nucleotide sequence of the fiber gene of ChAd4 (SEQ ID NO: 7).

Figures 12 A and B provides the nucleotide sequence of the fiber gene of ChAd5 (SEQ ID NO: 8).

Figures 13 A and B provides the nucleotide sequence of the fiber gene of ChAd7 (SEQ ID NO: 9).

Figures 14 A and B provides the nucleotide sequence of the fiber gene of ChAd9 (SEQ ID NO: 10).

Figures 15 A and B provides the nucleotide sequence of the fiber gene of ChAd10 (SEQ ID NO: 11).

Figures 16 A and B provides the nucleotide sequence of the fiber gene of ChAd11 (SEQ ID NO: 12).

Figures 17 A and B provides the nucleotide sequence of the fiber gene of ChAd16 (SEQ ID NO: 13).

Figures 18 A and B provides the nucleotide sequence of the fiber gene of ChAd17 (SEQ ID NO: 14).

Figures 19 A and B provides the nucleotide sequence of the fiber gene of ChAd19 (SEQ ID NO: 15).

Figures 20A-20D provides a comparison of the amino acid sequences of the fiber proteins of C1, ChAd11, ChAd20, ChAd17, ChAd3, ChAd19, PAN6, ChAd5, ChAd6, ChAd7, PAN5, PAN7, ChAd9, ChAd10, ChAd4, CV68 and ChAd16.

Figures 21A-21C provides the nucleotide sequence of the hexon gene of ChAd20 (SEQ ID NO: 16).

Figures 22A-22C provides the nucleotide sequence of the hexon gene of ChAd4 (SEQ ID NO: 17).

Figures 23A-23C provides the nucleotide sequence of the hexon gene of ChAd5 (SEQ ID NO: 18).

Figures 24A-24C provides the nucleotide sequence of the hexon gene of ChAd7 (SEQ ID NO: 19).

Figures 25A-25C provides the nucleotide sequence of the hexon gene of ChAd9 (SEQ ID NO: 20).

Figures 26A-26C provides the nucleotide sequence of the hexon gene of ChAd10 (SEQ ID NO: 21).

Figures 27A-27C provides the nucleotide sequence of the hexon gene of ChAd11 (SEQ ID NO: 22).

Figures 28A-28C provides the nucleotide sequence of the hexon gene of ChAd16 (SEQ ID NO: 23).

Figures 29A-29C provides the nucleotide sequence of the hexon gene of ChAd17 (SEQ ID NO: 24).

Figures 30A-30C provides the nucleotide sequence of the hexon gene of ChAd19 (SEQ ID NO: 25).

Figures 31A-31M provides a comparison of the amino acid sequences of the hexon proteins of : hAd12, hAd3, hAd7, hAd11, hAd21, hAd34, hAd35, C1, hAd1, hAd2, hAd5, ChAd3, ChAd11, ChAd17, ChAd19, ChAd20, hAd48, ChAd4, ChAd5 ChAd7, ChAd16, Pan6, hAd4, hAd16, ChAd6, ChAd9, ChAd10, C68, Pan5, Pan7, hAd41 and hAd40.

Figure 32 provides a listing of the oligonucleotide sequences (SEQ ID NOS: 26-46) disclosed herein.

Figure 33 is a graphic representation of the immunization break-point of ChAd vectors belonging to different serotype subgroups (i.e., subgroups C, E and D). The lowest dose eliciting a measurable immune response was determined by performing titration experiments in mice immunized with gag-expressing ChAd3, ChAd11, ChAd20, CV33, CV68, ChAd6, ChAd9, ChAd10, CV32, ChAd4, ChAd7 and ChAd16 vectors.

Figure 34 provides a graphic representation of a CEA-specific T cell response elicited in rhesus macaques immunized sequentially with a human adenoviral vector (MRKAd5 RhCEA) followed by a chimpanzee adenoviral vector (CV33 RhCEA) after 12 week interval. The immune responses were evaluated by IFN- γ ELISPOT assay, and the data illustrate the number of spot-forming cells (SFC) per million peripheral blood mononuclear cells (PBMC)

following incubation in the absence (DMSO) and presence of rhesus CEA C and D peptide pools.

Figure 35 provides a phylogenetic tree of human and chimpanzee adenoviruses of deduced from a multiple sequence alignment of full-length hexon peptide sequences using PAUPSEARCH (Wisconsin Package Version 10.3, Accelrys Inc.) and visualized and manipulated with TREEVIEW.

Figure 36 is a graphic representation of immunization results obtained in response to the administration of ChAd3 and hAd5 gag vectors to mice which were pre-exposed to hAd5. Cell-mediated immunity was evaluated 3 weeks post-immunization by IFN- γ ELISPOT using purified splenocytes.

Figure 37 is a graphic representation of kinetics of anti-CEA CMI elicited in human CEA transgenic mice immunized with ChAd3hCEA and Ad5hCEA. CMI was evaluated by ICS of PBMC stimulated with CEA peptide pool. The results are expressed as % of IFN γ ⁺ CD8⁺/total PBMC.

DETAILED DESCRIPTION OF THE INVENTION

As used throughout the specification and appended claims, the following definitions and abbreviations apply:

The term "cassette" refers to a nucleic acid molecule which comprises at least one nucleic acid sequence that is to be expressed, along with its transcription and translational control sequences. Changing the cassette, will cause the vector into which is incorporated to direct the expression of different sequence or combination of sequences. In the context of the present invention, the nucleic acid sequences present in the cassette will usually encode an immunogen. Because of the restriction sites engineered to be present at the 5' and 3' ends, the cassette can be easily inserted, removed or replaced with another cassette.

The term "*cis*-acting element" refers to nucleotide sequences which regulate genes to which they are attached. *Cis*-acting elements present in DNA regulate transcription, and those transcribed into mRNA can regulate RNA processing, turnover and protein synthesis.

The term "vector" refers to some means by which DNA fragments can be introduced into a host organism or host tissue. There are various types of vectors including plasmid, virus (including adenovirus), bacteriophages and cosmids.

The term "promoter" refers to a recognition site on a DNA strand to which an RNA polymerase binds. The promoter forms an initiation complex with RNA polymerase to

initiate and drive transcriptional activity. The complex can be modified by activating sequences such as enhancers, or inhibiting sequences such as silencers.

The term "pharmaceutically effective amount" refers to an amount of recombinant adenovirus that is effective in a particular route of administration to transduce host cells and provide sufficient levels of transgene expression to elicit an immune response.

The term "replication-competent" recombinant adenovirus (AdV) refers to an adenovirus with intact or functional essential early genes (i.e., E1A, E1B, E2A, E2B and E4). Wild type adenoviruses are replication competent.

The term "replication-defective" recombinant AdV refers to an adenovirus that has been rendered to be incapable of replication because it has been engineered to have at least a functional deletion, or a complete removal of, a gene product that is essential for viral replication. The recombinant chimpanzee adenoviral vectors of the invention are replication-defective.

The term "mammalian" refers to any mammal, including a human being.

The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome. (e.g., about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, e.g., the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, e.g. of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

In general, adenoviral constructs, gene constructs are named by reference to the genes contained therein. For example, "pChAd3 ΔE1gag" refers to a plasmid construct which comprises a ChAd3 chimpanzee adenoviral genome deleted of the E1 region. In this plasmid, the E1 region is replaced by an immunogen expression cassette comprising an HIV gag gene

under the control of a human CMV promoter followed by a bovine growth hormone polyadenylation signal. Similarly, pCV33DE1-E3 NSmut, refers to a second plasmid construct disclosed herein which comprises a CV33 chimpanzee adenoviral genome, deleted of the E1 and E3 regions, which is replaced by an immunogen expression cassette comprising HCV non-structural genes under the control a human CMV promoter followed by a bovine growth hormone polyadenylation signal.

The abbreviation "Ag" refers to an antigen.

As used throughout the specification and in the appended claims, the singular forms "a," "an," and "the" include the plural reference unless the context clearly dictates otherwise.

Adenoviruses (Ads) are nonenveloped, icosahedral viruses that have been identified in several avian and mammalian hosts. Human Ads (hAd) belong to the Mastadenovirus genus which includes all known human and many Ads of animal (e.g., bovine, porcine, canine, murine, equine, simian and ovine) origin. Human adenoviruses are divided into six subgroups (A-F) based on a number of biological, chemical, immunological and structural criteria which include hemagglutination properties of rat and rhesus monkey erythrocytes, DNA homology, restriction enzyme cleavage patterns, percentage G+C content and oncogenicity (Straus, 1984, In *The Adenoviruses*, ed. H. Ginsberg, pps. 451-498, New York: Plenus Press, and Horwitz, 1990 In *Virology*, eds. B.N. Fields and D.M. Knipe, pps. 1679-1721). To date, 51 distinct serotypes have been recognized and grouped into subgroups on the basis of their hemagglutination properties and biophysical and biochemical criteria.

The adenoviral virion has an icosahedral symmetry and, depending on the serotype, a diameter of 60-90 nm. The icosahedral capsid consists three major proteins, hexon (II), penton base (III) and a knobbed fibre (IV) as well as a number of minor proteins (i.e., VI, VIII, IX, IIIa and IVa2) (W.C. Russel, *J. Gen. Virol.*, 81: 2573-2604 (2000)). One aspect of the preexisting immunity that is observed in humans is humoral immunity, which can result in the production and persistence of antibodies that are specific for viral proteins. The humoral response elicited by adenovirus is mainly directed against the major structural proteins: hexon, penton and fiber.

Published reports have established that titers comprising antibodies against multiple serotypes are common (Dambrosio, E. (1982) *J. Hyg. (London)* 89: 209-219) and that a substantial portion of the preexisting titers have neutralizing activity. Neutralizing immunity to adenovirus is type specific, and infection with a particular serotype of adenovirus confers immunity only to that serotype. Several reports have suggested that antibodies directed towards the hexon are the strongest and the most neutralizing (Toogood, C.I.A., Crompton, J. and Hay

R.T. (1992) *J.Gen. Virol.* 73, 1429-1435). Therefore, it is reasonable to assume that the epitopes responsible for type-specific neutralization are located within seven hypervariable regions identified by alignment of the hexon sequences deriving from different serotypes. (Crawford-Miksza, L and D.P.Schnurr. (1996) *J.Virol.* 70:1836-1844).

5 A direct correlation between the presence of type-specific neutralizing antibodies and the inability to elicit an immune response with a vector based on the same serotype has been established by different methods including the passive transfer of immune sera from treated to naïve animals. Generally speaking, preexisting humoral immunity for a specific viral serotype reduces the therapeutic efficacy of the vector administration. Moreover, the administration of a
10 vector based on a specific viral serotype elicits an immune-response against the vector that prevents the re-administration of the same serotype.

In a particular embodiment, the invention provides a method of circumventing the adverse effects associated with the consequences of preexisting immunity to common serotypes of hAds. More specifically, the invention contemplates the use of chimpanzee adenoviral vectors
15 characterized by a serotype that does not circulate in humans. Accordingly, the invention provides adenoviral (Chad) vectors which lack neutralizing B-cell epitopes that cross react with those of common human serotypes as a vaccine carrier.

Although it has been reported that adenoviral-specific cell mediated immunity (CMI) can be cross-reactive, vaccination studies based on repeated injections of multiple
20 serotypes demonstrated a higher efficiency than immunization schedules based on a single vector. These experiments further demonstrate that the main limitation of a vector administration for vaccine purposes is the humoral pre-existing immunity against the vector. Potential solutions to the problems associated with the use of a human adenovirus as a vaccine carrier include the administration of a higher dose of an adenovirus (e.g., a subgroup C serotype) that is predicted to
25 encounter a preexisting immune response, and the use of vectors based on rare human serotypes. However, the use of higher doses of vaccine increases the cost of the vaccine and risk of undesirable side effects and the results of preclinical testing suggest that human alternate serotypes are less immunogenic than hAd5 and hAd6.

In an attempt to avoid the problems of host humoral and cellular immune
30 responses against the adenoviral backbone elements of the vector, and to minimize the risk of using human adenovirus-derived vector stocks that may be contaminated with replication-competent adenoviruses (RCA), several nonhuman adenoviruses have been characterized and developed as vaccine carriers (Soudois, C. *et al* (2000) *J. Virology*, 74:10639-10649; Farina, S.F. *et al* (2001) *J. Virology*, 75:11603-11613; Cohen, C.J. *et al* (2002) *J. Gen. Virology*, 83:151-
35 155.) The premise underlying the use of nonhuman adenoviral sequences to circumvent the

problems associated with preexisting immunity is based on the observation that neutralizing antibodies to common human adenovirus serotypes are unlikely to cross-neutralize nonhuman viruses. However, the incompatibility of viral and cellular factors imposes a practical limitation on the vast majority of alternative vector systems (bovine, ovine, canine) which are characterized by the disadvantage of having to be propagated in non-human cell lines.

Wilson *et al.* have published a report describing the characterization of a replication-defective vector based on chimpanzee adenovirus type 68 (CV68) C68, which was originally isolated from a mesenteric lymph node of a chimpanzee (Basnight, M., *et al.* (1971) *Am. J. Epidemiol.* 94:166-171.), CV68 was fully sequenced and found to be similar in overall structure to human adenoviruses (Farina, S. F. *et al.*, *J. Virol.* 75(23): 11603-11613 (2001). The genome of the virus is 36,521 base pairs in length and has been described as being most similar to subgroup E of human adenoviruses, with 90% identity to most human Ad4 open reading frames that have been sequenced. The CV68 ITRs are 130 base pairs in length, and all of the major adenoviral early and late genes are present. CV68 is characterized by a serotype that does not circulate in humans and which lacks neutralizing B cell epitopes that cross-react with those of common human serotypes. Although Chimpanzee adenoviruses are similar to human adenoviruses crossreactive neutralizing immunity against chimpanzee serotypes has not been documented in humans (Farina, S. F. *et al.* *J. Virol.* (2001) 75(23):11603-13).

The recombinant vectors derived from CV68 are described as being sufficiently similar to human serotypes to support transduction of cells expressing the coxsackievirus and adenovirus receptor (Cohen, C. *et al.*, *J. Gen. Virol.* 83: 151-155 (2002). Significantly, CV68 is characterized by a sufficient level of similarity to human adenoviruses to support its replication 293 cells which harbor E1 from human adenovirus type 5 (Farina, S. F. *et al.*, *J. Virol.* 75(23): 11603-11613 (2001). Furthermore, based on the observation that the flanking sequences of the human serotype 5 E1 are nonhomologous with those of the CV68-derived vector sequences, it is predicted that homologous recombination will not occur. Thus, it has been predicted that there is a low likelihood that CV68-derived vaccine stocks will be contaminated with RCA.

The same group of investigators subsequently reported the use of CV68-derived adenoviral sequences as a vaccine carrier for induction of antibodies to the rabies virus glycoprotein in mice. A replication-defective version of CV68 was created by replacing the E1A and E1B genes with a minigene cassette. Mice immunized with an E1-deletion-containing adenoviral recombinant (AdC68rab.gp) comprising a transgene product encoding the rabies virus glycoprotein developed protective immunity to rabies virus and remained resistant to challenge with an otherwise lethal dose of rabies virus (Xiang, Z *et al.*, *J. Virol.* 76(5): 2667-2675 (2002). A second CV68 construct expressing a codon-optimized, truncated form of gag of

HIV-1 was recently reported to induce a vigorous gag-specific CD8⁺ T cell response in mice. The vaccine-induced response was shown to provide protection to challenge with a vaccinia gag recombinant virus (Fitzgerald, J. C. *et al.*, *J. Immunol.* 170: 1416-1422 (2003). Experimental vaccination of mice preimmunized to human adenovirus serotype 5 with CV68gag or Ad5gag vectors demonstrated a more pronounced reduction of gag-specific T cells and protection against viral challenge elicited by Ad5 than by CV68 vaccine. The reduction in efficacy of C68gag vaccine was attributed to a cross-reactivity of Ad5-specific CD8⁺ T cells (*Id.*).

Considered together this data suggests that simian-derived replication-defective adenoviral vectors may be more suitable for use as human vaccine carriers than vectors based on common human serotypes. As shown herein, the results of experiments in which mice that were strongly immunized against human Ad5 (Figure 36) can be immunized with ChAd3-gag adenoviral vectors indicate the preexisting anti-human Ad5 immunity did not reduce the gag-specific CMI response elicited by the ChAd vectors. These results are consistent with the conclusion that human Ad5 cross-reactive B and T-cell epitopes are not present in ChAd3- or ChAd6 vectors.

Generally speaking, the adenoviral genome is very well characterized and despite the existence of several distinct serotypes, there is some general conservation in the overall organization of the adenoviral genome with specific functions being similarly positioned. The nucleotide sequences of the chimpanzee adenoviruses C1 and CV68 disclosed by Wilson *et al.*, and the location of the E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 genes of each virus are provided in U.S. Patent No. 6,083,716 (Chimpanzee Adenovirus Vectors), and PCT published application WO 03/000851 (Methods for Rapid Screening of Bacterial Transformants and Novel Simion Adenoviral Proteins), the teachings of which are incorporated herein by reference.

Each extremity of the adenoviral genome comprises a sequence known as an inverted terminal repeat (ITRs), which is necessary for viral replication. The virus also comprises a virus-encoded protease, which is necessary for processing some of the structural proteins required to produce infectious virions. The structure of the adenoviral genome is described on the basis of the order in which the viral genes are expressed following host cell transduction. More specifically, the viral genes are referred to as early (E) or late (L) genes according to whether transcription occurs prior to or after onset of DNA replication. In the early phase of transduction, the E1, E2, E3 and E4 genes of adenovirus are expressed to prepare the host cell for viral replication. The virus can be rendered replication defective by deletion of the essential early-region 1(E1) of the viral genome. Brody *et al.*, 1994 *Ann N Y Acad Sci.*, 716:90-101. During the late phase, expression of the late genes L1-L5, which encode the structural

components of the virus particles is switched on. All of the late genes are under the control of a single promoter and encode proteins including the penton (L2), the hexon (L3), the 100 kDa scaffolding protein (L4), and the fiber protein (L5), which form the new virus particle into which the adenoviral DNA becomes encapsidated. Depending on the serotype of the virus, 10,000-100,000 progeny adenovirus particles can be generated in a single host cell. Ultimately, the adenoviral replication process causes lysis of the cells.

The replication-defective adenoviral vectors disclosed herein were constructed by deletion of specific nucleotide sequences from the disclosed chimpanzee nucleic acid sequences and insertion of sequences derived other DNA sequences that are useful for transgene insertion, expression or other genetic manipulations. Accordingly, the recombinant chimpanzee adenoviruses described herein may contain adenoviral sequences derived from one or more chimpanzee adenoviruses, or sequences from a chimpanzee adenovirus and from a human adenovirus. Suitable polynucleotide sequences can be produced recombinantly, synthetically or isolated from natural sources. Adenoviral sequences suitable for use in particular aspects of the invention include sequences which lack neutralizing B-cell epitopes that are cross-reactive with common human serotypes.

At a minimum, the recombinant chimpanzee adenovirus (e.g., vector) of the invention contain the chimpanzee adenovirus *cis*-acting elements necessary for replication and virion encapsidation, in combination with at least one immunogen expression cassette. Typically, the *cis*-acting elements flank the expression cassette which comprises a transgene that encodes at least one antigen. More specifically, the vectors of the invention contain the requisite *cis*-acting 5' inverted terminal repeat (ITR) sequences of the adenoviruses (which function as origins of replication), 3' ITR sequences, packaging/enhancer domains, and a nucleotide sequence encoding a heterologous molecule. Regardless of whether the recombinant vector comprises only the minimal adenoviral sequences or an entire adenoviral genome with only functional deletions in particular genes (e.g., the E1 and/or E3 or E4 regions), the vectors of the invention comprise a chimpanzee adenovirus capsid.

Generally, speaking the adenoviral vectors disclosed herein comprise a replication-defective adenoviral genome, wherein the adenoviral genome does not have a functional E1 gene, and an immunogen expression cassette which comprises: a) a nucleic acid encoding at least one immunogen against which an immune response is desired; and b) a heterologous (i.e., with respect to the adenoviral sequence) promoter operatively linked to the nucleic acid sequence encoding the immunogen(s); and a transcription terminator.

More specifically, the invention provides replication-defective vectors that consist of a recombinant adenoviral genome that is devoid of at least one early gene selected from the

group consisting of E1, E2, E3, and E4. In one embodiment, a replication-defective vector is prepared by replacing, or disrupting, the E1 gene of one of the adenoviral isolates disclosed herein (e.g., ChAd3, ChAd6, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19 or ChAd20) with an immunogen expression cassette. For example, a vector can be prepared by deleting/disrupting the E1 gene of ChAd 3 (SEQ ID NO:1) or ChAd6 (SEQ ID NOS: 2). Alternatively, a replication-defective vector can be prepared from any one of the other adenovirus isolates disclosed herein, including ChAd3, ChAd6, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19 or ChAd20. In other embodiments, replication-defective vectors of the invention comprises an adenoviral genome derived from one of the ChAds disclosed herein that has been optionally engineered to lack a functional E3 gene. It is to be understood that the chimpanzee adenoviral sequences disclosed herein can be rendered replication-defective by either completely removing an early gene or by rendering the gene inoperative or nonfunctional.

It is to be understood that the invention encompasses vectors that are characterized as having modifications, such as a "functional deletion" which destroys the ability of the adenovirus to express one or more selected gene products. The phrase "functional deletion" as used herein broadly encompasses modifications that have the effect of rendering a particular gene product nonfunctional. Generally speaking, functional deletions take the form of a partial or total deletion of an adenoviral gene. However, one of skill in the art will readily acknowledge that other manipulations, including but not limited to making a modification which introduces a frame shift mutation, will also achieve a functional deletion. For example, the recombinant chimpanzee adenoviral vectors of the invention can be rendered replication-defective by introducing a modification that is designed to interfere with, or to functionally delete, the ability of the virus to express adenoviral E1A and/or E1B.

It is well-known that replication-defective adenoviral vectors can be obtained by introducing a modification that is designed to interfere with, or to functionally delete the expression of one or more genes from the group of E2 genes. More in detail, a replication-defective vector can be constructed by inactivating the polymerase gene, or the pre-terminal protein gene or the DNA binding protein gene. Moreover deletion or inactivation of genes expressed by E4 region is an alternative strategy to construct replication-defective chimp Ad vectors. Early gene deletion or inactivation can be combined in order to produce more attenuated vectors. Alternatively, replication-defective ChAd vectors can also comprise additional modifications in other viral genes, such as the late genes L1 through L5. In addition, novel adenoviral vaccine carriers can be generated by combining hexon and fiber genes obtained from different serotypes. The utilization of a hexon and fiber gene shuffling strategy will also allow an

investigator to change the biological properties of a ChAd and facilitate the production of vectors with a different tropism or with new serological characteristics.

It is to be understood that the present invention encompasses recombinant adenoviral vectors comprising deletions of entire genes or portions thereof which effectively
5 destroy the biological activity of the modified gene either alone or in any combination. For example, recombinant simian adenoviruses can be constructed which have a functional deletion of the genes expressed by E4 region, although as shown herein it may be desirable to introduce the heterologous Ad5 E4 sequence into the vector in combination with the functional deletion of an E1 gene. Alternatively, the function of the adenoviral delayed early E3 gene may be
10 eliminated; however because the function of E3 is not necessary for the production of a recombinant adenoviral particle it is not necessary to replace this gene product in order to produce a recombinant that is capable of packaging a virus useful in the invention.

In one embodiment of this invention, the replication- defective adenoviral vector used is a chimpanzee subgroup C adenovirus containing deletions in E1 and optionally in E3.
15 For example, for ChAd3, a suitable E1 deletion/disruption can be introduced in the region from bp 460 to bp 3542 (with reference to SEQ ID NO: 1). For ChAd6, a suitable E1 deletion/disruption can be introduced in the region from bp 457 to bp 3425 (with reference to SEQ ID NO: 2). For CV32, the E1 deletion is preferably from bp 456 to bp 3416 (with reference to SEQ ID NO: 3); for CV33, the E1 deletion is preferably from bp 456 to bp 3425 (with
20 reference to SEQ ID NO: 4) and for CV23, the E1 deletion is preferably from bp 456 to bp 3415 (with reference to SEQ ID NO: 5). E3 deletions for CV32 and CV33 are preferably from bp 27446 to bp 31911 (with reference to SEQ ID NO: 3); from bp 27146 to bp 31609 (with reference to SEQ ID NO: 4) respectively. Those of skill in the art can easily determine the equivalent sequences for other chimpanzee isolates based on sequence homologies and multiple
25 sequence alignments.

One of skill in the art will readily acknowledge that in order to construct an E1-deleted adenoviral vector a number of decisions must be made regarding the structure of the vector backbone and the composition of the nucleic acid sequence comprising the transgene. For example, an investigator must determine if the size of the E1 deletion will accommodate the size
30 of the transgene. If not, then additional deletions will have to be introduced into the backbone of the vector.

The nucleic acid sequence embodying the transgene can be a gene, or a functional part of a gene and will typically exist in the form of an expression cassette. Typically a gene expression cassette includes: (a) nucleic acid encoding a protein or antigen of interest; (b) a
35 heterologous promoter operatively linked to the nucleic acid encoding the protein; and (c) a

transcription termination signal. The nucleic acid can be DNA and/or RNA, can be double or single stranded. The nucleic acid can be codon-optimized for expression in the desired host (*e.g.*, a mammalian host).

5 Decisions must also be made regarding the site within the backbone where the transgene will be introduced and the orientation of the transgene. More specifically, the transgene can be inserted in an E1 parallel (transcribed 5' to 3') or anti-parallel (transcribed in a 3' to 5' direction relative to the vector backbone) orientation. In addition, appropriate transcriptional regulatory elements that are capable of directing expression of the transgene in the mammalian host cells that the vector is being prepared for use as a vaccine carrier in need to be
10 identified and operatively linked to the transgene. "Operatively linked" sequences include both expression control sequences that are contiguous with the nucleic acid sequences that they regulate and regulatory sequences that act in *trans*, or at a distance to control the regulated nucleic acid sequence.

15 Regulatory sequences include: appropriate expression control sequences, such as transcription initiation, termination, enhancer and promoter sequences; efficient RNA processing signals, such as splicing and polyadenylation signals; sequences that enhance translation efficiency (*e.g.*, Kozak consensus sequences); sequences that enhance protein stability, and optionally sequences that promote protein secretion. Selection of these and other common vector elements are conventional and many suitable sequences are well known to those
20 of skill in the art (see, *e.g.*, Sambrook *et al.*, and references cited therein at, for example, pages 3.18-3.26 and 16.17-16.27 and Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989).

25 In specific embodiments, the promoter is a heterologous promoter (*i.e.*, with respect to the adenovirus sequences) which is recognized by an eukaryotic RNA polymerase. In a preferred embodiment, the promoter is a "strong" or "efficient" promoter. An example of a strong promoter is the immediate early human cytomegalovirus promoter (Chapman *et al.*, 1991 *Nucl. Acids Res* 19:3979-3986, which is incorporated by reference). The human CMV promoter can be used without (CMV) or with the intron A sequence (CMV-intA), although those skilled in the art will recognize that any of a number of other known promoters, such as the strong
30 immunoglobulin, or other eukaryotic gene promoters may be used, including the EF1 alpha promoter, the murine CMV promoter, Rous sarcoma virus (RSV) promoter, SV40 early/late promoters and the beta-actin promoter.

35 Further examples of promoters that can be used in the present invention are the strong immunoglobulin promoter, the EF1 alpha promoter, the murine CMV promoter, the Rous Sarcoma Virus promoter, the SV40 early/late promoters and the beta actin promoter, albeit those

of skill in the art can appreciate that any promoter capable of effecting expression in the intended host can be used in accordance with the methods of the present invention. The promoter may comprise a regulatable sequence such as the Tet operator sequence. Sequences such as these that offer the potential for regulation of transcription and expression are useful in instances where repression of gene transcription is sought.

Suitable gene expression cassettes will also comprise a transcription termination sequence. A preferred transcriptional terminator is the bovine growth hormone terminator. The promoter/transcription termination combination of CMVintA-BGH terminator is particularly preferred although other promoter/terminator combinations may also be used. As shown herein, the bovine growth hormone termination/polyadenylation signal (bGHpA) or short synthetic polyA signal (SPA) of 50 nucleotides in length defined as follows:

AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTT-GGTTTTTTGTGTG (SEQ ID NO:26). Generally speaking, exemplify suitable termination sequences. The polyA signal is inserted following the nucleic acid sequence which comprises the transgene and before the 3' adenovirus ITR sequence.

The recombinant adenoviral vectors described herein may contain adenoviral sequences derived from one or more strain of adeonvirus. Suitable sequences may be obtained from natural sources, produced recombinantly, synthetically, or by other genetic engineering or chemical methods. In a particular embodiment, the recombinant chimpanzee adenovirus is a chimeric recombinant comprising non-chimpanzee adenoviral polynucleotide sequences. Suitable non-chimpanzee adenoviral sequences can be obtained from human adenoviral strains. For example, the native E4 region can be replaced by hAd5 E4 (Ad5 nt 32816 to nt 35619) or by Ad5E4orf6 (Ad5 nt 33193 to nt 34077) (Ad5 GenBank Accession No: M73260).

Generally speaking, the immunogen (antigenic molecule) delivered by the recombinant adenoviral vector of the invention comprises a polypeptide, protein, or enzyme product which is encoded by a transgene in combination with a nucleotide sequence which provides the necessary regulatory sequences to direct transcription and/or translation of the encoded product in a host cell. The composition of the transgene depends upon the intended use of the vector. For example, if the immunogenic composition is being designed to elicit an antibody response or a cell-mediated immune response in a mammalian host which is specific for an infectious agent, then it is appropriate to utilize a nucleic acid sequence encoding at least one immunogenic product that is predicted to confer pathogen-specific immunity to the recipient. Alternatively, if the composition is being prepared for use as a cancer vaccine, a suitable transgene may comprise an immunogenic portion of a self-antigen, such as a TAA, which has been selected with the goal of eliciting a protective immune response of sufficient potency to

both break host tolerance to a particular TAA and to elicit a long-lived (e.g., memory) response that will be sufficient to prevent the initiation of cancer or to prevent tumor progression.

Accordingly, suitable immunogenic gene products may be obtained from a wide variety of pathogenic agents (such as, but not limited to viruses, parasites, bacteria and fungi) that infect mammalian hosts, or from a cancer or tumor cell. Although, the invention is illustrated herein with a particular set of test immunogens it is to be understood that the invention is not limited to the use of the antigens exemplified herein. More specifically, the invention contemplates the use of both heterologous and self-antigens as immunogens, including but not limited to TAAs.

In one embodiment, the invention provides an immunogenic composition (e.g., a vaccine) for inducing an immune response against antigens (i.e., immunogens) expressed by an infectious agent. For example, it is desirable to elicit an immune response against a virus infecting humans and/or non-human animal species. Examples of virus families against which a prophylactic and/or therapeutic immune response would be desirable include the *Picornaviridae* family which includes six different genera such as Aphthovirus, Cardiovirus, Enterovirus, Hepatovirus, Parechovirus, Rhinovirus. Examples of Picornavirus against which an immuneresponse would be desirable are: Foot-and-mouth disease viruses, Encephalomyocarditis viruses, Polioviruses, Coxsackieviruses, Human hepatitis A virus, Human parechoviruses, Rhinoviruses. *Caliciviridae* family includes different genera associated with epidemic gastroenteritis in humans caused by the Norwalk group of viruses and other syndromes in animals like the hemorrhagic disease in rabbits associated with rabbit hemorrhagic disease virus or respiratory disease in cats caused by feline calicivirus.

Another family of viruses, against which it may be desirable to elicit an immune response is the *Astroviridae* which comprises viruses isolated from humans as well as many different animal species. Human astroviruses are associated with gastroenteritis and young children diarrhea. Alternatively, it may be desirable to confer mammalian hosts with immunity to members of the *Togaviridae* family of viruses which comprises two genera: alphavirus and rubivirus. Alphaviruses are associated with human and veterinary diseases such as arthritis (i.e. Chikungunya virus, Sindbis virus) or encephalitis (i.e. Eastern Equine Encephalitis Virus, Western Equine Encephalitis Virus).

Rubella virus provides an alternative viral target against which is the only member of the Rubivirus genus is responsible for outbreaks of a mild exanthematic disease associated with fever and lymphadenopathy. Rubella virus infection is also associated with fetus abnormalities when acquired by mother during in early pregnancy. *Flaviviridae* is an other virus family consisting of three genera: the flaviviruses, the pestiviruses and the hepaciviruses that includes important human as well as animal pathogens. Many of the flavivirus genus members

are arthropod-borne human pathogens causing a variety of diseases including fever, encephalitis and hemorrhagic fevers. Dengue Fever Viruses, Yellow Fever Virus, Japanese Encephalitis Virus, West Nile Fever Virus, Tick-borne Encephalitis Virus are pathogens of major global concern or of regional (endemic) concern. Pestivirus genus includes animal pathogens of major economic importance such as Bovine Viral Diarrhea Virus, Classical Swine Fever Virus, Border Disease Virus. Hepatitis C Virus is the only member of the Hepacivirus genus responsible for acute and chronic hepatitis. HCV proteins expressed by a recombinant adenovirus can elicit a protective as well as therapeutic immune response limiting the consequences of a viral infection affecting 170 million people worldwide.

Alternatively, antigens derived from members of the *Coronaviridae* family can be expressed by recombinant adenovirus vectors in order to obtain protection against infection. Protection against the severe acute respiratory syndrome coronavirus (SARS-Co Virus) can be obtained by immunizing with one or more chimpanzee adenovirus chosen from the group including ChAd3, 4, 5, 6, 7, 9, 10, 11, 16, 17, 19, 20 expressing one or more SARS-CoV protein including without limitations nucleocapsid (N) protein, polymerase (P) protein, membrane (M) glycoprotein, spike (S) glycoprotein, small envelope (E) protein or any other polypeptide expressed by the virus. *Rhabdoviridae* family members including rabies virus can be target of recombinant vaccine expressing viral proteins.

Other possible targets include the *Filoviridae* family comprising Ebola-like viruses and Marburg-like viruses genera, that is responsible of outbreaks of severe hemorrhagic fever; the *Paramyxoviridae* family comprising some of the most prevalent virus known in humans like measles, respiratory syncytial, parainfluenza viruses and viruses of veterinary interest like Newcastle disease and rinderpest viruses; the *Orthomyxoviridae* family including Influenza A, B, C viruses; *Bunyaviridae* family mainly transmitted by arthropod to vertebrate hosts comprising important human pathogens like Rift valley fever, Sin Nombre, Hantaan, Puumala viruses; *Arenaviridae* family comprising Lymphocytic choriomeningitis, Lassa fever, Argentine Hemorrhagic fever, Bolivian Hemorrhagic fever viruses; *Bornaviridae* family comprising viruses causing central nervous system diseases mainly in horses and sheep; *Reoviridae* family including rotaviruses, the most important cause of severe diarrheal illness in infants and young children worldwide, orbiviruses that can affect both humans and other mammals (bluetongue, epizootic hemorrhagic disease viruses); *Retroviridae* family, a large group of viruses comprising important human pathogens like human immunodeficiency virus 1 and 2 (HIV-1 and HIV-2) and human t-cell leukemia virus type 1 and 2 (HTLV 1 and 2) as well as non-human lentivirus such as Maedi/Visna viruses affecting sheep and goats, Equine infectious anemia virus affecting horses, bovine immunodeficiency virus affecting cattle, feline immunodeficiency virus affecting

cats; Polyomaviridae family groups small DNA oncogenic viruses, prototype viruses are polyoma and SV40 infecting mouse and rhesus monkey respectively, (BK and JC viruses closely related to SV40 were isolated from human patients); *Papillomaviridae* family consists of a group of DNA viruses infecting higher vertebrates including humans generating warts and condylomas.

- 5 Papilloma viral infection is associated with the development of cancer in both humans and animals. Human papilloma viruses are associated with cervical cancer, vaginal cancer and skin cancer. The herpesviridae famils includes subfamilies in which are classified a number of important pathogens for humans and other mammals. Suitable sources of antigens can be but are not limited to herpes simplex viruses 1 and 2, varicella-zoster virus, Epstein-Barr virus,
- 10 Cytomegalovirus, human herpesviruses 6A,6B and 7, Kaposi's sarcoma-associated herpesvirus. Further suitable source of antigens are members of the Poxviridae family like Monkeypox virus, Molluscum contagiosum virus, smallpox virus; Hepatitis B virus, the prototype member of the hepadnaviridae family as well as other virus causing acute and/or chronic hepatitis like hepatitis delta virus, hepatitis E virus.
- 15 The adenoviral vectors of the present invention are also suitable for the preparation of immunogenic compositions designed to stimulate an immune response in humans or animals against protein expressed by non-viral pathogens including bacteria, fungi, parasites pathogens For example,the vectors disclosed herein can be used to prepare vaccines against, but not limited to: *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*,
- 20 *Vibrio cholerae*, *Clostridium tetani*, *Neisseria meningitis*, *Corynebacterium diphtheriae*, *Mycobacteria tuberculosis* and *leprae*, *Listeria monocytogenes*, and *Legionella pneumofila*. Examples of fungi and mammalian parasites for which it may be desirable to prepare prophylactic or therapeutic vaccines include: *Candida albicans*, *Aspergillus fumigatus*, *Histoplasma capsulatum*, *Plasmodium malariae*, *Leishmania major*, *Trypanosome cruzi* and
- 25 *brucei*, *Schistosoma haematobium* , *mansoni* and *japonicum*; *Entamoeba histolytica*, and numerous species of *Filaria* known to be responsible for human filariasis.

- Cancer typically involves the deregulation of genes that encode polypeptides which contribute to maintaining cell cycle or controlling cell proliferation (e.g., growth factors, oncogenes, receptors and tumor suppressors). The products of many of the genes implicated in
- 30 cancer are expressed on the surface of a wide variety of tumor cells. A variety of tumor antigens that may be recognized by T and B lymphocytes have been identified in human and animal cancer. The vast majority of human tumor-associated antigens (TAAs) that are suitable for use in an anticancer vaccine trial are described as "self-antigens" due to the fact that in addition to being expressed on tumor cells they also are expressed on normal tissue and/or during fetal

development. Immunotolerance of the target population to TAAs may explain why many cancer vaccines have proven to be ineffective.

5 Tumor antigens can be produced by oncogenic mutants of normal cellular genes altered proto-oncogenes or tumor suppressor genes such as Ras, p53 or Bcr-Abl protein are
10 examples of altered cellular proteins that can stimulate T/B cell response. Tumor antigens can be normal cellular proteins that are overexpresses in tumor cells (tyrosinase, GP100, MART are normally expressed at low levels in melanocytes and overexpressed in melanoma) or aberrantly expressed in tumor cells (MAGE, BAGE, GAGE expressed in melanomas and many carcinomas but normally expressed in the testis and placenta). Tumor antigens can be products of oncogenic
15 viruses: papillomavirus E6 and E7 proteins expressed by cervical carcinomas; EBV EBNA-1 protein produced by EBV+ lymphomas and nasopharyngeal carcinomas; SV40 T antigen in SV40 induced experimental tumors. Oncofetal antigens are expressed to high levels on cancer cells and in normal developing (fetal) tissues but not in adult tissues. Carcinoembryonic antigen (CEA) and alpha-fetoprotein (AFP) are examples of well characterized oncofetal antigens.

15 Recent evidence supports the existence of TAAs that are capable of eliciting an immune response, thus making this class of antigens suitable immunogens for vaccine therapy. However, as a class of antigens TAAs are notoriously poor immunogens and T cells that are highly specific for TAAs are either deleted or anergized during T-cell development. Accordingly, there is an expectation that the immune response of a tumor-bearing host to a
20 particular TAA will be extremely weak. Because of the inherent need to break host tolerance to a target TAA experimental clinical vaccine studies are particularly focused on developing immunization strategies that will enhance TAA-specific T-cell responses. Generally, speaking an effective cancer vaccine must both overcome immunotolerance and enhance host's immune response to a level that is preventative and/or protective. Anti-tumor effects in many experimental
25 vaccine studies have been correlated with T-cell responses to TAAs.

30 In an alternative embodiment, the invention contemplates an immunogenic composition (e.g., a cancer vaccine) which can be used to induce an immune response against tumor antigens. A suitable composition would contain a recombinant chimpanzee adenovirus comprising nucleic acid sequence encoding a tumor antigen and a physiologically acceptable carrier. In a particular embodiment, the coding sequence element of the cassette may encode a single immunogen, such as an immunogenic peptide sequence derived from a self-antigen, such as a tumor-associated antigen. In some embodiments, the nucleic acid sequence encoding the immunogen (i.e., the transgene) may be codon optimized for expression in a particular
35 mammalian species. In other embodiments, the coding sequence may encode more than one immunogen, such as one or more codon optimized tumor antigens. For example, a cancer

vaccine utilizing the disclosed adenoviral vectors may encode a combination of self-antigens such as: HER2/neu, CEA, Hepcam, PSA, PSMA, Telomerase, gp100, Melan-A/MART-1, Muc-1, NY-ESO-1, Survivin, Stromelysin 3, Tyrosinase, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, hKLP2, VEGF.

5 Development of an effective cancer vaccine requires the identification of a strategy that will elicit antigen-specific immunity in vaccinated patients and the generation of an immune response that will persist after active immunization has ended. The success of the strategy will depend on whether a measurable immune response directed against a target antigen will correlate with protection against cancer occurrence or relapse. The effector mechanisms of both cell-mediated immunity and humoral immunity have been show to kill tumor cells. However, data from experimental systems suggest that antigen-specific T cells represent the most powerful immunologic mechanism for the elimination of tumor cells. Recognition of tumor-specific antigens (e.g., TAAs) by effector T-cells is predicted to allow the TAA to function as a tumor-rejection antigen. Published studies suggest that stimulation of CD8⁺ and CD4⁺ helper T-cell responses are important for achieving optimal tumor clearance ((Greenberg, P. D. (1991) *Adv. Immunol.* 49: 281-355; Pardoll, D. M. *et al.* (1998) *Curr. Opin. Immunol.* 10: 588-94). Clinical response (i.e., efficacy) has been associated with increases in interferon γ -secreting cytotoxic T cells. The advent of assays, such as the ELISPOT assay used herein, to demonstrate the efficacy of the instant vaccine carriers, allows investigators to measure T-cell responses to vaccination regimens and thereby facilitates the development of cancer vaccines.

20 Cancer vaccines can be either prophylactic or therapeutic. The general assumption underlying the prophylactic use of cancer vaccines is that TAAs are extremely weak immunogens or functionally nonimmunogenic in tumor-bearing subjects. More specifically, in the field of cancer immunology, vaccines can be used as immunotherapy in patients afflicted with cancer. Accordingly, cancer vaccines can be designed to elicit an immune response that is that is directed against a TAA that is expressed by a pre-existing tumor or malignancy. Thus, in particular embodiments, therapeutic cancer vaccines are intended for use in tumor-bearing patients who have developed resistance to conventional regimens of treatment or who have a high probability of developing a recurrence following conventional treatment.

30 The high immunogenicity of adenoviruses, make adenoviral vectors particularly good candidates for use in the context of a vaccine carrier designed to break host tolerance to a self-antigen. The phenomenon of epitope or determinant spreading, which was first described in autoimmune diseases, has been associated with both MHC class I- and MHC class II-restricted responses and correlated to the development of HER-2/neu protein-specific T-cell immunity. Epitope spreading represents the generation of an immune response to a particular portion of an

immunogenic protein followed by the natural spread of immunity to other antigenic determinants present on the same protein. For example, Disis *et al.* observed epitope spreading in 84% of patients afflicted with HER-2/neu overexpressing malignancies who were administered vaccines comprising peptides derived from potential T-helper epitopes of the HER-2 protein mixed with granulocyte-macrophage colony stimulating factor (*J. Clin. Oncol.* (2002) 20(11): 2624-2632). Importantly, epitope spreading was correlated with the generation of a HER-2/neu protein domain response and suggests that immunization effectively circumvented immunologic tolerance.

TAAAs that are suitable for use in the disclosed adenoviral vectors and methods as a target for a cancer vaccine should possess a number of characteristics. For example, a target TAA must have a favorable expression profile, meaning that it should be preferentially expressed or overexpressed in the tumor or malignant tissue as compared with normal tissue. In addition, because TAAAs that play a role in tumorigenesis are more likely to be retained during the different stages of cancer progression, a suitable target TAA should also be preserved throughout tumor progression and metastases. Suitable target TAAAs should also be expressed homogeneously within the tumor. Third, suitable target TAAAs must not be subject to absolute immunologic tolerance. More specifically, there should be some evidence that T cells which can both recognize and respond to the TAA of interest have not been entirely deleted from the host's T-cell repertoire (Berinstein, N. L., *J. Clin. Oncol.* 29(8): 2197 (2002).

Carcinoembryonic antigen (CEA) has many characteristics which make it an attractive TAA for use as a target antigen for an anticancer vaccine. It is a member of the Ig superfamily which is characterized by a favorable expression pattern. It is expressed in more than 50% of all human cancers and has been implicated in the tumorigenesis process, which suggests that its expression may be selected and conserved throughout cancer progression. In addition, it has been established that immunologic tolerance to CEA is not absolute. Published studies establish that human T cells can recognize, become activated to, and lyse cancer cells that express CEA (Berinstein, N. L., *J. Clin. Oncol.* 29(8): 2197 (2002). For example, the immunization of patients with recombinant vaccinia virus expressing CEA, combined with subsequent peptide-based in vitro stimulations, generated CD8+ MHC-restricted CTLs capable of lysing autologous tumors (Tsang, K. Y. *et al. J. Natl. Cancer Inst.*, (1995) 87:982-990). Alternatively, immunization of colorectal carcinoma patients after surgery with recombinant CEA was reported to induce weak antibody and cellular responses to recombinant CEA (Samanci, A., *et al.* (1998) *Cancer Immunol. Immunother.* 47: 131-142.) Further, the administration of anti-CEA anti-idiotypic antibody to patients diagnosed with colorectal cancer generated anti-CEA antibodies and idiotypic-specific T-cell proliferation (Foon, L, A. *et al.*

(1995) *J. Clin. Invest.*, 96: 334-342). The literature also indicates that tolerance to CEA in cancer patients can be overcome with several different vaccination approaches (i.e., vaccination with recombinant CEA or recombinant orthopox or avipox-CEA viruses, administration of anti-idiotypic antibodies, pulsing dendritic cells with CEA agonist epitopes).

5 CEA is an oncofetal glycoprotein that is expressed in normal fetal colon and to a much lesser extent in normal colonic mucosa. It is also overexpressed in the vast majority of adenocarcinomas, particularly those of the colon, pancreas, breast, lung, rectum and stomach. Many colorectal cancers and some carcinomas produce high levels of CEA that are measurable in sera, which makes it one of the most widely used serological markers of malignancy, especially
10 in patients with colorectal cancer.

A second TAA which provides a suitable immunogen for use in the compositions and methods of the invention is product of the HER2/erb-2 (also called neu) proto-oncogene. Like, CEA, HER2/neu has a favorable expression pattern and is not subject to absolute tolerance. More specifically, low levels of expression of the HER2/neu transcript, and the 185 kD
15 polypeptide product, are detected in normal adult epithelial cells of various tissues, including the skin and breast, and tissues of the gastrointestinal, reproductive, and urinary tracts; higher levels of expression are detected in the corresponding fetal tissues during embryonic development (Press *et al.*, *Oncogene* 5: 953-962 (1990). Several lines of evidence suggest a link between the amplification of HER-2 and neoplastic transformation in human breast, lung, prostate, ovarian,
20 endometrial and colorectal tumors (Disis and Cheever, *Adv. Cancer Research* 71: 343-371(1997). Generally speaking, overexpression of HER2/neu correlates with a poor prognosis and a higher relapse rate for cancer patients (Slamon *et al.*, *Science* 244: 707-712 (1989). Thus, a vaccine specific for the HER-2/neu protein could have wide application and utility in the prevention of disease recurrence in many different human malignancies.

25 HER2/neu encodes a transmembrane glycoprotein possessing intrinsic tyrosine kinase activity and displaying extensive homology to the epidermal growth factor (EGF) receptor (Akiyama, T *et al.*, (1986) *Science* 232: 1644-1646). One of the first clinical studies which utilized HER2 as target for cancer immunotherapy employed the HER-2-specific monoclonal antibody Herceptin for the treatment of breast cancer (Goldenberg MM (1999) *Clin. Ther.* 21:
30 309-318). This led to subsequent efforts which focused on the use of HER-2 as a target for the T-cell arm of the immune system to elicit effective antitumor responses, including the use of recombinant fusion proteins comprising HER-2 domains to activate autologous antigen presenting cells. Published reports establish that numerous cancer patients afflicted with neu-expressing mammary and ovarian cancers mount immune responses (e.g., produce antigen-
35 specific antibody and T-cells) against the protein product of the HER2/neu oncogene.

Assembly of the recombinant adenoviral sequences, transgene and other vector elements into various intermediate plasmids and shuttle vectors, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques as described in standard textbooks that are well known to those of skill in the art (Sambrook *et al*,
 5 Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989). Such techniques include, but are not limited to conventional cDNA cloning techniques, use of overlapping oligonucleotide sequences derived from the adenoviral genome, homologous recombination, polymerase chain reaction, standard transfection techniques, plaquing of viruses in agar overlay and other related methodologies.

10 To assist in preparation of polynucleotides in prokaryotic cells, a plasmid version of the adenovirus vector is often prepared (adenovirus pre-plasmid). The adenovirus pre-plasmid contains an adenoviral portion and a plasmid portion. The adenoviral portion is essentially the same as the adenoviral portion contained in the adenoviral vectors of the invention (containing adenoviral sequences with non-functional or deleted E1 and optionally E3 regions)
 15 and an immunogen expression cassette, flanked by convenient restriction sites.

The plasmid portion of the adenovirus pre-plasmid often contains an antibiotic resistance marker under transcriptional control of a prokaryotic promoter so that expression of the antibiotic does not occur in eukaryotic cells. Ampicillin resistance genes, neomycin resistance genes and other pharmaceutically acceptable antibiotic resistance markers may be
 20 used. To aid in the high level production of the polynucleotide by fermentation in prokaryotic organisms, it is advantageous for the adenovirus pre-plasmid to contain a prokaryotic origin of replication and be of high copy number. A number of commercially available prokaryotic cloning vectors provide these benefits. It is desirable to remove non-essential DNA sequences. It is also desirable that the vectors not be able to replicate in eukaryotic cells. This minimizes the
 25 risk of integration of polynucleotide vaccine sequences into the recipients' genome. Tissue-specific promoters or enhancers may be used whenever it is desirable to limit expression of the polynucleotide to a particular tissue type.

Adenovirus pre-plasmids (plasmids comprising the genome of the replication-defective adenovirus with desired deletions and insertions) can be generated by homologous
 30 recombination using adenovirus backbones DNA and an appropriate shuttle vector (designed to target-in specific deletions and incorporate desired restriction sites into the resultant plasmid). Shuttle vectors of use in this process can be generated using general methods widely understood and appreciated in the art, *e.g.*, PCR of the adenoviral terminal ends taking into account the desired deletions, and the sequential cloning of the respective segments into an appropriate
 35 cloning plasmid. The adenoviral pre-plasmid can then be digested and transfected into the

complementing cell line via calcium phosphate co-precipitation or other suitable means. Virus replication and amplification then occurs, a phenomenon made evident by notable cytopathic effect. Infected cells and media are then harvested after viral replication is complete (generally, 7-10 days post-transfection).

5 Generally speaking, following the construction and assembly of the desired adenovirus pre-plasmids, adenovirus pre-plasmids are rescued into virus by transfecting an adenoviral E1-expressing human cell line. Complementation between the packaging cell line and the viral genes of the vector permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the production of recombinant
10 adenoviruses. The resulting viruses may be isolated and purified by any of a variety of methods known to those of skill in the art for use in the methods of the invention.

 It will be readily apparent to those of skill in the art that when one or more selected deletions of chimpanzee adenoviral genes are introduced into a viral vector, the function of the deleted gene product can be supplied during the production process by sequences present
15 in the production cell line. Thus, the function of the manipulated genes can be provided by a permanently transformed cell line that is characterized by some or all of the adenoviral functions which are required for packaging but which are not functional in the vector (e.g., any of E1A, E1B, E2A, E2B E4). Alternatively, the requisite adenoviral functions can be provided to a suitable packaging cell line by infecting or transiently transfecting a suitable cell with a construct
20 comprising the requisite gene to provide the function.

 Accordingly, the present invention also provides a method of producing chimpanzee adenoviral vectors in E1-expressing human cell lines. More specifically, the disclosed vectors can be propagated in an E1 complementing cell lines, including the known cell lines 293 and PER.C6™. Both these cell lines express the adenoviral E1 gene product.
25 PER.C6™ is described in WO 97/00326, published January 3, 1997, which is hereby incorporated by reference. It is a primary human retinoblast cell line transduced with an E1 gene segment that complements the production of replication deficient first generation adenoviruses, but is designed to prevent generation of replication competent adenovirus by homologous recombination. 293 cells are described in Graham *et al* (1977) *J. Gen. Virol* 36:59-72, which is
30 also hereby incorporated by reference. One of skill in the art will recognize the term "first generation adenovirus" refers to a replication deficient adenovirus which has either a non-functional or deleted E1 region, and optionally a non-functional or deleted E3 region.

 Batches of replication-defective adenoviral vectors that are intended for use as a vaccine composition in a clinical trial should be proven to be free of RCA (Fallaux, F.J. *et al*
35 (1998) *Humm Gene Therapy*, 9:1909-1917). In practice, this is a labor intensive process which

requires establishing and utilizing an expensive screening program. One of skill in the art will acknowledge that a high frequency of RCA generation not only results in a high failure rate for the batches produced, but also severely limits scale-up efforts. Elimination of sequence homology between the nucleotide sequence of the vector and the adenoviral sequences present in the genome of the helper production/packaging cell line should eliminate the possibility of producing batches of vector that are contaminated with RCAs produced by homologous recombination.

Typically, recombinant replication-defective adenoviral vectors are propagated in cell lines that provide E1 gene products *in trans*. Supplementation of the essential E1 gene products *in trans* is very effective when the vectors are from the same or a very similar serotype. For example, it is well-known that E1-deleted (i.e. Δ E1) group C serotype (Ad2 and Ad5) vectors, can be propagated in 293 or PER.C6 cells which contain and express the Ad5 E1 region. However, it has been observed that Ad5 E1 sequences present in the 293 and PER.C6 production cells may not always fully complement the replication of non-group C serotypes. Accordingly, E1-deleted serotypes outside of subgroup C, for example those from subgroups A, B, D, E, and F may replicate with a lower efficiency respect to the corresponding wt virus or may not replicate at all in 293 or PER.C6 cells. This may be due to the inability of the Ad5 (group C) E1B 55K gene product to establish a functional interaction with the E4 orf6 gene product of the non-group C serotypes.

The decrease in replication efficiency in cells expressing Ad5 E1 is variable considering vectors of different subgroups. While Δ E1 vectors deriving from subgroup D and E adenovirus can be rescued and propagated in 293 and Per.C6TM cells with variable efficiency, the propagation Δ E1 vectors of subgroup B is completely impaired (Vogels R, *et. al.* (2003) Aug. Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: efficient human cell infection and bypass of preexisting adenovirus immunity. *J Virol.*; 77 (15):8263-71).

Although the interaction between Ad5 E1b 55k and vector-expressing E4 orf6 protein is conserved within members of the same subgroup, it may be not sufficiently stable when E4 orf6 protein of a non-C serotype is expressed. This inefficient or unstable formation of E1B-55K/E4-orf6 complex lead to an absent of reduced propagation of the Δ E1 vector. Accordingly, it has been empirically determined that in order to successfully and efficiently rescue recombinant adenovirus of groupB serotypes, a cell line expressing the E1 region of the serotype of interest may need to be generated. In cells expressing Ad5E1 like 293 or Per.C6TM, the expression can be limited to E1b 55K protein. Alternatively, a suitable Ad5E1-expressing cell lines could be modified to express the entire Ad5 E4 region (or E4 orf6 only) in addition to

Ad5E1. The generation of cell lines expressing both Ad5 E1 and orf6 are useful in complementing alternative adenovirus serotypes; *see, e.g.,* Abrahamsen *et al.*, 1997 *J. Virol.* 8946-8951. The incorporation of E4 (orf6) into Ad5 complementing cell lines, is known, as is the generation of serotype-specific cell lines providing a serotype-specific E1 gene product(s) *in trans*. Alternatively, the efficiency of non-group C vector propagation may be improved by modification of the viral backbone by substituting the native E4 region with Ad5 orf6. Similar results can be achieved by substituting the only the native orf6 with orf6 deriving from Ad5 or other subgroup C viruses (Ad1, Ad2, Ad6). U.S. Patent No. 5,849,561 discloses complementation of an E1-deleted non-group C adenovirus vector in an Ad5-E1 complementing cell line which also expresses portions of the Ad5-E4 gene.

U.S. Patent No. 6,127,175, issued to Vigne, *et al.*, discloses a stably transfected mammalian cell line which expresses a portion of the E4 region of adenovirus, preferably orf6/orf6/7. Such a cell line is useful for complementation of recombinant Ad genomes deficient in the E4 region.

Compositions, including vaccine compositions, comprising the disclosed adenoviral vectors are an important aspect of the present invention. These compositions can be administered to mammalian hosts, preferably human hosts, in either a prophylactic or therapeutic setting. Potential hosts/vaccinees include but are not limited to primates and especially humans and non-human primates, and include any non-human mammal of commercial or domestic veterinary importance. Compositions comprising recombinant chimpanzee adenoviral vectors may be administered alone or in combination with other viral- or non-viral-based DNA/protein vaccines. They also may be administered as part of a broader treatment regimen.

In a particular embodiment of the invention, the disclosed vectors may be used in an immunization protocol designed to break host tolerance to a self-antigen or a tumor-associated antigen. The identification of a number of TAA has enabled the development of active vaccination approaches for the therapy of cancer. Both cell surface antigens and intracellular antigens that are processed and presented provide useful targets. Generally speaking, the disclosed method of breaking host tolerance to a self-antigen comprises: (a) stimulating an antigen-specific response to a self-antigen by administering a first vaccine composition comprising a first ChAd vector or a plasmid vector carrying a nucleotide sequence encoding the self-antigen against which an antigen-specific immune response is desired, and (b) sustaining and expanding the immune response of (a) by administering a second vaccine composition comprising a recombinant ChAd vector of a different serotype containing at least a functional deletion of its genomic E1 gene, and in the site of the E1 gene, a sequence comprising a promoter capable of directing the expression of DNA encoding the same self-antigen delivered in the

priming step, whereby the host mounts an immune response which has the effect of breaking tolerance to the self-antigen.

Accordingly, a skilled artisan can utilize this disclosure to design several different immunization protocols that may be suitable for use to break host tolerance. For example, it may be possible to utilize a protocol in which the first, or priming immunization comprises plasmid DNA which encodes a particular self-antigen, such as a TAA, and any subsequent immunizations comprise a ChAd vector. Plasmid DNA sequences comprising nucleotide sequences that encode self-antigens, may be delivered intramuscularly, with or without electrostimulation, in one or more injections. For example, an immunization protocol based on multiple (e.g., 3 or 4 or 5) intramuscular injections of plasmid DNA encoding a TAA via electroporation followed by one or more intramuscular injections of a ChAd vector comprising a transgene encoding the same TAA is encompassed by the general method disclosed and claimed herein.

Alternatively, a suitable protocol to break tolerance could involve one or more priming immunizations with a ChAd or hAd vector comprising a transgene encoding a self antigen, followed by one or more boosting immunizations with either the same, or a different ChAd vector that is known to be non cross-reactive with the vector used for the priming immunization(s). For example, an immunization protocol using ChAd3 for priming and ChAd6 for boosting, or ChAd3 for priming followed by ChAd6 and ChAd9 for boosting could be used to break host tolerance. In particular embodiments, the invention contemplates the use of self-antigens comprising at least one tumor associated antigen selected from the group consisting of: HER2/neu, CEA, EpCAM, PSA, PSMA, Telomerase, gp100, Melan-A/MART-1, Muc-1, NY-ESO-1, Survivin, Stromelysin 3, Tyrosinase, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, hKLP2, VEGF. In a particular embodiment, the invention provides a method for inducing an immune response (e.g., humoral or cell-mediated) to a tumor-associated antigen which is specific for a selected malignancy by delivering a recombinant chimpanzee adenovirus encoding the TAA to a mammal afflicted with cancer. In a preferred embodiment of this aspect of the invention the elicited immune response constitutes an immune response characterized by the production of antigen-specific CD4+ and CD8+ T cells.

The immunogenic compositions of the invention can be administered to mammalian hosts, preferably human hosts, in either a prophylactic or therapeutic setting. Potential hosts/vaccinees include but are not limited to primates and especially humans and non-human primates, and include any non-human mammal of commercial or domestic veterinary importance. Compositions comprising recombinant chimpanzee adenoviral vectors may be

administered alone or in combination with other viral- or non-viral-based DNA/protein vaccines. They also may be administered as part of a broader treatment regimen. Suitable compositions, for use in the methods of the invention may comprise the recombinant viral vectors of the invention in combination with physiologically acceptable components, such as buffer, normal saline or phosphate buffered saline, sucrose, other salts and polysorbate. It does not cause tissue irritation upon intramuscular injection. It is preferably frozen until use. Optionally, a vaccine composition of the invention may be formulated to contain other components, such as but not limited to, an adjuvant, a stabilizer, a pH adjusting agent, or a preservative. Such components are well known to those of skill in the art.

It is envisioned that the recombinant chimpanzee adenoviruses of the invention will be administered to human or veterinary hosts in an "effective amount," that is an amount of recombinant virus which is effective in a chosen route of administration to transduce host cells and provide sufficient levels of expression of the transgene to invoke an immune response which confers a therapeutic benefit or protective immunity to the recipient/vaccine.

The amount of viral particles in the vaccine composition to be introduced into a vaccine recipient will depend on the strength of the transcriptional and translational promoters used and on the immunogenicity of the expressed gene product. In general, an immunologically or prophylactically effective dose of 1×10^7 to 1×10^{12} particles (i.e., 1×10^7 , 2×10^7 , 3×10^7 , 5×10^7 , 1×10^8 , 2×10^8 , 3×10^8 , 5×10^8 or 1×10^9 , 2×10^9 , 3×10^9 , 5×10^9) and preferably about 1×10^{10} to 1×10^{11} particles is administered directly into muscle tissue. Subcutaneous injection, intradermal introduction, impression through the skin, and other modes of administration such as intraperitoneal, intravenous, or inhalation delivery are also contemplated.

The recombinant chimpanzee adenoviral vectors of the present invention may be administered alone, as part of a mixed modality prime/boost vaccination regimen or in a vaccination regimen based on combination of multiple injections of different vector serotypes. Typically, a priming dose(s) comprising at least one immunogen is administered to a mammalian host in need of an effective immune response to a particular pathogen or self-antigen. This dose effectively primes the immune response so that, upon subsequent identification of the antigen(s), the host is capable of immediately mounting an enhanced or boosted immune response to the immunogen. A mixed modality vaccination scheme which utilized alternative formulations for the priming and boosting can result in an enhanced immune response. Prime-boost administrations typically involve priming the subject (by viral vector, plasmid, protein, *etc.*) at least one time, allowing a predetermined length of time to pass, and then boosting (by viral vector, plasmid, protein, *etc.*). Multiple immunizations, typically 1-4, are usually employed, although more may be used. The length of time between priming and boost

may typically vary from about four months to a year, albeit other time frames may be used as one of ordinary skill in the art will appreciate. Multiple injection of each vector can be administered within approximately a 2 weeks time frame, before neutralizing immunity becomes evident.

In some embodiments of this invention, a vaccine is given more than one
5 administration of adenovirus vaccine vector, and it may be given in a regiment accompanied by the administration of a plasmid vaccine. Suitable plasmid vaccines for use in combination with the vectors disclosed herein comprise a plasmid encoding at least one immunogen against which a primed or boosted immune response is desired, in combination with a heterologous promoter, which is capable of directing expression of the nucleic acid sequences encoding the
10 immunogen(s), operably linked to the immunogen coding sequence, and a transcription terminator sequence.

For example, a dosing regimen which utilizes multiple injection of different serotypes of recombinant replication-defective chimpanzee adenoviral vectors can be used. Alternatively, an individual may be given a first dose (i.e., a priming dose) of a plasmid vaccine,
15 and a second dose (i.e., a boosting dose) which comprises a replication-defective recombinant chimpanzee adenoviral vector which comprises a coding sequence for the same immunogen that was delivered in the plasmid vaccine. Alternatively, the individual may be given a first dose of a human adenovirus vaccine vector encoding at least one immunogen, followed by a second dose comprising a replication-defective recombinant chimpanzee adenoviral vector disclosed herein,
20 which comprises a coding sequence for the same immunogen that was delivered in the priming dose. In a second alternative embodiment a vaccine composition comprising a vector of the invention may be administered first, followed by the administration of a plasmid vaccine. In any of these embodiments, an individual may be given multiple doses of the same immunogen, in either viral vector or plasmid form. There may be a predetermined minimum amount of time
25 separating the administrations.

In addition to a single protein or antigen of interest being delivered by the recombinant, replication-defective chimpanzee adenovirus vectors of the present invention, two or more proteins or antigens can be delivered either via separate vehicles or delivered *via* the same vehicle. Multiple genes/functional equivalents may be ligated into a proper shuttle plasmid
30 for generation of a adenovirus pre-plasmid comprising multiple open reading frames. Open reading frames for the multiple genes/functional equivalents can be operatively linked to distinct promoters and transcription termination sequences.

As shown herein, suitable immunization regimens can employ different adenoviral serotypes. One example of such a protocol would be a priming dose(s) comprising a
35 recombinant adenoviral vector of a first serotype, for example a ChAd3 or ChAd6 followed by a

boosting dose comprising a recombinant chimpanzee adenoviral vector of a second serotype. In an alternative embodiment, the priming dose can comprise a mixture of separate adenoviral vehicles each comprising a gene encoding for a different protein/antigen. In such a case, the boosting dose would also comprise a mixture of vectors each comprising a gene encoding a separate protein/antigen, provided that the boosting dose(s) administers recombinant viral vectors comprising genetic material encoding for the same or similar set of antigens that were delivered in the priming dose(s). These multiple gene/vector administration modalities can further be combined. It is further within the scope of the present invention to embark on combined modality regimes which include multiple but distinct components from a specific antigen.

Use of recombinant vectors derived from chimpanzee adenoviruses that are not neutralized by preexisting immunity directed against the viral elements of human vector offers an alternative to the use of human Ad vectors as vaccine carriers. Because adenoviruses are highly immunogenicity, adenoviral vectors are particularly good candidates for use in the context of a vaccine carrier designed to break host tolerance to a self-antigen. Furthermore, the ability to propagate the chimp viruses in human cells, particularly in the Per.C6TM cell line, with an efficiency comparable to human viruses, offers considerable advantages both from a regulatory point of view and for the large scale production of therapeutics or vaccines. Accordingly, the instant invention provides a collection of chimpanzee adenoviral sequences, vectors and plasmids that allow the preparation of recombinant virus which may be used, alone or in combination, as a vaccine carrier for genetic vaccination.

All publications mentioned herein are incorporated by reference for the purpose of describing and disclosing methodologies and materials that might be used in connection with the present invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

Having described preferred embodiments of the invention with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments, and that various changes and modifications may be effected therein by one skilled in the art without departing from the scope or spirit of the invention as defined in the appended claims.

The following examples illustrate, but do not limit the invention.

EXAMPLE 1 ISOLATION, CLONING, SEQUENCING AND CHARACTERIZATION OF ChAds

Chimpanzee Adenovirus Isolation

Stool specimens were collected in viral transport medium (VTM; Microtest M4-R Multi-Microbe Transport Medium, Remel Inc.) then frozen or frozen directly at -70°C at NIRC (New Iberia Research Center 4401 W. Admiral Doyle Drive New Iberia, LA 70560). The specimens were kept frozen at $< -70^{\circ}\text{C}$ until they were processed for inoculation into cell cultures. At that time, the specimens were thawed and then vortexed in excess of chilled viral transport medium. After the specimens had dissociated into suspensions, they were centrifuged for 10 min at 1500-1800 rpm. The supernatants were filtered through 0.8 and 0.2 μm syringe filters in series and then the filtered material was inoculated into cell cultures (200-250 μL into shell vials and 250-300 μL into tube cultures). Each processed specimen was inoculated into tube cultures and shell vial cultures seeded with 293 cells or A549 cells.

Control (positive and negative) cultures were prepared each time a set of samples was inoculated. Once all of the shell vials in a set-up had been inoculated, they were centrifuged at room temperature for 60 ± 10 min at 2000 rpm ($900 \times g$). The vials were removed from the centrifuge immediately after the rotor stopped spinning to prevent heat damage in the cultures. After centrifugation, the inocula were aspirated from the shell vials, using a fresh sterile pasteur pipet in each vial to prevent cross-contamination. The cultures were washed three times using 1.0-mL fresh culture medium for each wash. Fresh medium (1.0 mL) was pipetted into each vial after the third wash and the shell vials were placed in an incubator at $35-37^{\circ}\text{C}$ for three to four days (approx. 96 hr).

At the end of the culture period, the supernatants were aspirated from the cultures and the cell layer in each vial was washed twice with Immunofluorescence Assay (IFA) Buffer using approximately 1.0 mL buffer with each wash. The cells were fixed by adding 1.0 mL refrigerated acetone to each vial (10 min at $2-8^{\circ}\text{C}$). Acetone-cleaned slides were labeled with the specimen identification number(s) associated with the shell vial coverslips. The shell vial coverslips were processed for fluorescence labeling of Adenovirus-infected cells using a primary mouse anti-adenovirus antibody [MAB8052, Chemicon]. The slides are evaluated with the aid of a fluorescence microscope. Each preparation was scanned using the 10X objective noting the extent of immunofluorescence coverage across the well (1+ to 4+). The presence or absence of specific immunofluorescence was confirmed using the 40X objective. Tube cultures were inoculated in the same sequence as described for the shell vials (e.g., negative control first, followed by clinical specimens and positive controls). The inocula were allowed to adsorb for

60-120 min at 36-38°C. After the adsorption period, the specimens/controls were aspirated from the tubes and replaced by fresh culture medium.

Three to four days post-inoculation, and once a week thereafter, the media was aspirated from the culture tubes and replaced with 1.5 mL fresh media. Culture tubes were visually monitored for CPE at least every other day for at least 21 days after inoculation. Cultures inoculated with chimp specimens were compared against the controls and rated by observing the CPE extent. Cultures showing no CPE were passed to fresh tube cultures after 14 days; culture tubes that were negative for CPE after 21 days were considered negative. Culture tubes with 3-4+ CPE were vortexed for 10 seconds. The cells were scraped from the wall of the tube using a sterile 1.0 mL serological pipet and suspended in the culture supernatant. After labeling a 5 mL snap cap tube with the specimen identification number and date and stored at -70°C. 500 µL of the cell suspension was transferred from the culture tube into the snap cap tube and stored for up to one day at 2-8°C until it was processed using an indirect immunofluorescent antibody technique to detect adenovirus (equivalent to procedure for staining shell vials).

Chimpanzee Adenovirus Amplification

Wild type chimp adenoviruses CV32, CV33, CV23 and CV68 purchased from the ATCC (ATCC Accession Numbers: CV32, VR-592; CV-33, VR-593;) or from Esoterix Inc. Austin, Texas and original isolates were propagated as follows by using the human E1-expressing cell line PER.C6™ or 293. Briefly, cells were cultivated in Dulbecco's Modified Eagles Medium (DMEM; GibcoBRL, Life Technologies) supplemented with 10% Fetal Bovine Serum (FBS GibcoBRL, Life Technologies), 1% Penicillin-Streptomycin, 2mM Glutamine and 10mM MgCl₂ (Per.C6™). Adenovirus infection was carried out in DMEM supplemented with 5% Horse Serum (GibcoBRL, Life Technologies). Infected cells and medium were collected when 100% of the cells exhibited virus-induced cytopathic effect (CPE) and lysed by three cycles of freezing and thawing.

All wild type chimp adenoviral (CV) stocks were cloned by infecting 293 cells seeded in 96-well plates, after the first passage of amplification. The virus cloning was performed by limiting dilution of the cell lysate obtained at the first passage of the virus amplification. 5 isolated clones were picked up and serially propagated. After 3-4 serial passaging of amplification, a large-scale preparation of adenovirus was performed on cells planted on 5 two-layer cell-factories (NUNC) (200 millions of cells/cell factory). Purified viral particles were obtained from cell lysate by two ultra-centrifugation steps on cesium chloride density gradients.

Sequencing of Viral Genomic DNA

Genomic DNA was isolated from 3×10^{12} pp of purified virus preparation by digestion with Proteinase K (0.5 mg/ml) in 1% SDS-TEN (2 hrs at 55°C). After a Phenol-Chloroform extraction and Ethanol precipitation, the genomic DNA was resuspended in water and submitted for genomic sequencing.

For full length Ad genome sequencing, the purified viral DNA was nebulized to produce randomly sheared fragments. The DNA fragments were blunt-ended with the klenow fragment of E.coli DNA polymerase and polynucleotide kinase. The blunt end fragment were run on a low melting point agarose gel to purify the fragments in the size range of 1-3 kb and cloned into the SmaI site of pUC19 vector to create a shotgun library. The ligations were used to transform competent XL1-Blue MRF'. Positive colonies were identified by white/blue screening on LB agar containing X-gal and IPTG. Three to four 96-well block of plasmid DNA were isolated from the library and sequenced with pUC forward and reverse primers. All sequencing reads were screened for quality and vector sequence using the Phred-Phrap software package. The reads that passed the screening were assembled into contigs. Primers were designed to directly sequence the adenoviral DNA for closing the gaps and determine the DNA sequence of both ends.

Complete viral genome sequencing was obtained for selected viruses including ChAd3 (SEQ ID NO: 1), ChAd6 (SEQ ID NO: 2) and CV32 (SEQ ID NO:3), CV33 (SEQ ID NO: 4), and CV23 (SEQ ID NO:5). Table 1 provides data summarizing the percentage of identity between the nucleotide sequences of ChAd3, ChAd6, Pan5 (CV23), Pan6 (CV32), Pan7 (CV33), C1 and C68 adenoviral genomes. Alignments were calculated using the ALIGN program as part of the FASTA package version 2 (William R. Penson, University of Virginia; Myers & Miller, CABIOS 1989, 4:11-17).

Table 1. Percentage of Nucleotide Sequence Identity Between Chimpanzee Adenovirus Genomes

	ChAd3	ChAd6	Pan5	Pan6	Pan7	C1	C68
ChAd3	100	68.1	68.5	68.2	68.3	64.2	68.0
ChAd6		100	95.5	94.5	95.5	73.6	91.4
Pan5			100	94.9	96.7	73.9	92.7
Pan6				100	95.1	73.6	91.3
Pan7					100	73.8	93.0
C1						100	74.3
C68							100

5 To characterize the new adenoviral isolates (e.g., ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17 and ChAd19) the nucleotide sequence of the hexon and fiber genes were also determined by primer walking. Fiber gene: SEQ ID NOS: 6-15: (SEQ ID NO: 6, ChAd20); SEQ ID NO: 7, ChAd4); SEQ ID NO: 8, ChAd5); SEQ ID NO: 9, ChAd7); SEQ ID NO: 10, ChAd9); SEQ ID NO: 11, ChAd10); SEQ ID NO: 12, ChAd11); SEQ
10 ID NO: 13, ChAd16) SEQ ID NO: 14, ChAd17) and SEQ ID NO: 15, ChAd19). Figures 20A-20D provide a comparison of the amino acid sequences of the fiber proteins of the ChAd isolates disclosed and claimed herein.

The hexon gene sequences are set forth in SEQ ID NOS: 16-25: (SEQ ID NO: 16, ChAd20); SEQ ID NO: 17, ChAd4); SEQ ID NO: 18, ChAd5); SEQ ID NO: 19, ChAd7); SEQ
15 ID NO: 20, ChAd9); SEQ ID NO: 21, ChAd10); SEQ ID NO: 22, ChAd11); SEQ ID NO: 23, ChAd16); SEQ ID NO: 24, ChAd17) and SEQ ID NO: 25, ChAd19). Figures 31A-31M provide a comparison of the amino acid sequences of the hexon proteins of the ChAd isolates disclosed and claimed herein.

20 Chimpanzee Adenovirus Classification

Classification of the different chimp adenoviral strains follows the already proposed classification of human adenovirus serotypes into 6 subgroups (Horowitz, MS (1990) Adenoviridae and their replication. In Virology B.N. Fields and D.M. Knipe, eds (Raven Press, New York) pp.1679-1740) and it was obtained by amino acid and nucleotide sequence alignment
25 by using Align X program (Informax, Inc).

An initial classification of the new isolates was obtained by looking at the restriction pattern of the viral genome with different restriction endonucleases and by sequence

analysis of the hypervariable region 7 (HVR7) of the hexon gene. To this end two primers were designed on the highly conserved regions flanking HVR7: TGTCCTACCARCTCTTGCTTGA (SEQ ID NO.45) and GTGGAARGGCACGTAGCG (SEQ ID NO.46). The HVR7 was amplified by PCR using purified viral DNA or crude 293 lysate as template and then sequenced.

5 Based on HVR7 sequence analysis we classified the new isolated viruses into the subgroups (A-F) proposed for human Ad viruses (Horowitz, MS (1990) Adenoviridae and their replication. In Virology B.N. Fields and D.M. Knipe, eds (Raven Press, New York) pp.1679-1740).

10 The phylogenetic tree presented in Figure 35 was obtained by alignment of human and chimp adenovirus hexon amino acid sequences. The results are consistent with the initial classification based on nucleotide sequence alignment limited to hexon HVR7 by using Align X program (Informax, Inc). The tree was deduced from a multiple sequence alignment of full-length hexon peptide sequences using a PAUPSEARCH (Wisconsin Package Version 10.3, Accelrys Inc.) and visualized and manipulated with TREEVIEW. Bootstrap confidence analysis was performed using the PAUPSEARCH program as implemented in the Wisconsin Package.

15 For each of the alignments the program was run on 1000 replicates using "Heuristic Search" as search criterion and Maximum Parsimony as the optimality criterion and confidence values reported were taken from a 50% majority-rule consensus.

20 **EXAMPLE 2 ChAd SHUTTLE VECTOR AND EXPRESSION VECTOR CONSTRUCTION AND RESCUE**

Vector Construction and Rescue

Genomic viral DNA was cloned into a standard plasmid vector by homologous recombination with an appropriate shuttle vector containing viral DNA sequences derived from both left and right end of viral genome (Figure 2). As described more fully below, the sequence

25 homology observed between viruses classified in the same serotype subgroup was exploited to develop group-specific shuttle vectors. Genomic viral DNA of Chimp adenovirus classified into subgroup D and E resulted to be sufficiently homologous to allow the construction of a common shuttle vector in order to clone viruses belonging to both subgroups.

30 **Construction of a Subgroup D/E Shuttle Vector**

The ChAd6 viral genome was fully sequenced (SEQ ID NO: 2) and the information obtained was used to construct a shuttle vector to facilitate cloning by homologous recombination of subgroup D and E chimpanzee adenovirus.

Construction of the ChAd6 shuttle vector, referred to herein as

pARS ChAd6-3 is described in Figure 1. Figure 32 provides a list of the oligonucleotide sequences (SEQ ID NOS: 26-40 and SEQ ID NOS: 45-46) used in the cloning experiments described herein. Briefly, 457 bp deriving from the left end of ChAd6 DNA were amplified by PCR with the oligonucleotides 5'-ATGGAA

- 5 TTCGTTTAAACCATCATCAATAATATACCTC-3 (SEQ ID NO: 27) and 5'-CGCTGGCACTCAAGAGTGGCCTC-3' (SEQ ID NO: 28) digested with EcoRI and SnaBI and cloned into pNEBAd35-2 cut EcoRI-SnaBI, generating pNEBChAd6-LI. The right ChAd6 ITR (bp 36222 to bp 36648) was amplified by PCR using the oligonucleotides: 5'-ATGAAGCTTGTTTAAACCCATCATCAATAATATACCT-3' (SEQ ID NO: 29) and 5'-ATCTAGACAGCGTCCATAGCTTACCG-3' (SEQ ID NO: 30) digested with restriction enzymes HindIII and XbaI and cloned into pNEBChAd6-LI HindIII-XbaI digested thus generating pNEBChAd6-RLI. Finally, the DNA fragment corresponding to nucleotides 3426-3813 of the ChAd6 genomic DNA sequence was amplified with the oligonucleotides: 5' ATGCTACGTAGCGATCGCGTGAGTAGTGTGTTGGGGGTGGGTGGG-3' (SEQ ID NO: 31) and 5'-TAGGCGCGCCGCTTCTCCTCGTTCAGGCTGGCG-3' (SEQ ID NO: 32), digested with SnaBI and AscI then ligated with SnaBI-AscI digested pNEBChAd6-RLI thus generating pNEBChAd6-RLIdE1.

- 20 To improve the efficiency of recombination and plasmid propagation in DH5a *E.coli* strain, the 1306 bp fragment containing both left and right ITRs of ChAd6 as well as pIX gene fragment was excised by PmeI digestion from pNEBChAd6-RLIdE1 and transferred to a different plasmid vector obtained by PCR amplification with the oligonucleotides 5'-GATCTAGTTAGTTTAAACGAATTCGGATCTGC GACGCG-3' (SEQ ID NO: 33) and 5' TTCGATCATGTTTAAACGAA ATTAAGAATTCGGATCC-3' (SEQ ID NO: 34) from pMRKAd5SEAP. This final ligation step generated the ChAd6 shuttle vector pARSchAd6-3.

Construction of a Subgroup C Shuttle Vector

- 30 The ChAd3 viral genome was fully sequenced (SEQ ID NO: 1) and the information obtained was used to construct a shuttle vector to facilitate cloning by homologous recombination of subgroup C chimpanzee adenovirus.

- 35 Briefly, the shuttle vector used to clone subgroup C chimp adenovirus, referred to herein as pChAd3EGFP was constructed as follows: a ChAd3 DNA fragment (nt 3542-4105) containing pIX coding region was amplified by PCR with the oligonucleotides 5'-

TATTCTGCGATCGCTGAGGTGGGTGAGTGGGCG-3' (SEQ ID NO: 35) and 5'-TAGGCGCGCCCTTAAACGGCATTTGTGGGAG-3' (SEQ ID NO: 36) digested with SgfI-AscI then cloned into pARSCV32-3 digested with SgfI-AscI, generating pARS-ChAd3D. ChAd3 right end (nt 37320-37441) was amplified by PCR with oligonucleotides 5'-CGTCTAGAAGACCCGAGTCTTACCAGT-3' (SEQ ID NO: 37) and 5'-CGGGATCCGTTTAAACCATCATCAATAATATACCTTATT-3' (SEQ ID NO: 38) digested with XbaI and BamHI then ligated to pARS-ChAd3D restricted with XbaI and BamHI, generating pARS-ChAd3RD. ChAd3 viral DNA left end (nt 1-460) was amplified by PCR with oligonucleotides 5'-ATGGAATTTCGTTTAAACCATCATCAATAATATACCTT-3' (SEQ ID NO: 39) and 5'-ATGACGCGATCGCTGATATCCTATAATAATAAAACGCAGACTTTG-3' (SEQ ID NO: 40) digested with EcoRI and SgfI then cloned pARS-ChAd3RD digested with EcoRI and SgfI, thus generating pARS-ChAd3RLD. The viral DNA cassette was also designed to contain restriction enzyme sites (PmeI) located at the end of both ITR's so that digestion will release viral DNA from plasmid DNA.

Construction of Δ E1 Chimp Adenoviral Vectors

Subgroup C: Subgroup C chimp adenovirus vectors were constructed by homologous recombination in *E. coli* strain BJ5183. BJ5183 cells were co-transformed with pChAd3EGFP shuttle vector digested with BstEII and Bst1107I and ChAd3, ChAd11, ChAd19 and ChAd20 purified viral DNA. Homologous recombination between pIX genes, right ITR DNA sequences present at the ends of linearized pChAd3EGFP and viral genomic DNA allowed its insertion in the plasmid vector, deleting at the same time the E1 region that was substituted by EGFP expression cassette. Expression cassettes based on human cytomegalovirus (HCMV) promoter and bovine growth hormone polyadenylation signal (Bgh polyA) were constructed to express secreted alkaline phosphatase (SEAP), EGFP, HIV gag, HCV NS region (as described in fig.3 for ChAd6 shuttle vectors) as well as tumor-associated antigens like CEA and HER2/neu from human and Rhesus monkey origin.

Subgroups D and E: In order to construct Δ E1 vectors based on subgroup D and E chimp adenovirus, the shuttle vector pARS ChAd6-3 was digested with AscI and co-transformed into *E. coli* strain BJ5183 with CV32, CV33, CV68, ChAd4, ChAd5, ChAd6, ChAd7, ChAd9, ChAd10 and ChAd16 purified viral DNA. Homologous recombination between DNA sequences from pIX genes and right ITR present at the ends of linearized pARS ChAd6-3 and viral genomic DNA allowed its insertion in the plasmid vector, deleting at the same time the E1 region (figures 2 and 4).

Expression cassettes based on human cytomegalovirus (HCMV) promoter and bovine growth hormone poly-adenylation signal (Bgh polyA) were constructed to express secreted alkaline phosphatase (SEAP), EGFP, HIV gag, HCV NS genes (Figure 3) as well as tumor-associated antigens like CEA and HER2/neu of human and Rhesus monkey origin. All the expression cassette were inserted into the single SnaBI site of pARS ChAd6-3 vector to be transferred by homologous recombination into the Δ E1 adenovirus pre-plasmids as described in figure 4.

Rescue and amplification of Δ E1 Vectors

5×10^6 PER.C6TM cells planted on 6cm cell culture dishes were transfected with 10 micrograms of cloned viral vector released from plasmid sequences by endonuclease digestion. DNA transfection was performed using Lipofectamine (Invitrogen). Transfected cells and culture medium were collected 5-10 days post-transfection and lysed by freeze-thaw. Rescued vectors were then amplified by serial passaging on 293 or PER.C6TM cells. A large-scale amplification was performed by infecting cells planted on 5-10 cell-factories (NUNC, Inc.) on a total of $1-2 \times 10^9$ cells. A purified vector preparation was obtained on cesium chloride gradient by two ultra-centrifuge runs, dialyzed against PBS containing 10% glycerol and stored at -70°C in aliquots.

EXAMPLE 3 NEUTRALIZATION STUDIES

Neutralization assays were carried out in order to evaluate the prevalence in human sera of neutralizing antibodies against the chimpanzee adenoviruses disclosed herein. The assay evaluated the effects of serum preincubation on the ability of chimp adenoviruses carrying the gene for secreted alkaline phosphatase (SEAP) to transduce human 293 cells. The neutralization titer is defined as the dilution of serum giving a 50% reduction of the SEAP activity observed in the positive control with the virus alone.

From 2×10^6 to 1.5×10^7 physical particles of CV33-SEAP, CV32-SEAP and ChAd3-SEAP vector were diluted in 100 μl of complete medium and added to an equal volume of human or chimp serum diluted in complete medium. Each serum samples was tested at various dilutions (five 4-fold increments starting from 1/18 dilution through 1:4608). Samples were pre-incubated for one hour at 37°C and then added to 293 cells seeded into 96-well plates (3×10^4 cells/well). The inoculum was removed after one hour of incubation, the cells were re-fed with fresh medium and, 24 hours later, 50 μl of medium was removed and the SEAP activity was measured by a chemiluminescent assay. The neutralization titer is defined as the dilution of serum giving a 50% reduction of the SEAP activity observed in the positive control with the virus alone. A panel of 100 human sera was tested for ChAd neutralization activity. In parallel the same panel was tested on Ad5 SEAP vector.

5 **Table 2. Prevalence of neutralizing antibodies against chimpanzee adenovirus**

	Virus			
titer	hAd5	CV32	CV33	ChAd3
< 200	77%	96%	100%	92%
> 200	33%	4%	0%	8%

10 The result provided in Table 2 indicates that a very low prevalence in human sera of neutralizing antibodies directed against vector derived from chimpanzee adenoviruses. Only four sera showed a titer over the threshold of 200 on CV32 vector while 8 showed a titer over 200 on ChAd3 SEAP vector. On the contrary, the panel of chimp sera examined showed a very high prevalence of anti-Chimp Ad immunity. These findings confirm that as expected, vectors based on chimp Ads have a very little chance to be neutralized in humans. Therefore they represents an ideal solution to the problem of the pre-existing anti-human Ad immunity that
15 limits the administration of viral vectors based on common human Ad serotypes such as Ad5.

MURINE IMMUNIZATION STUDIES

METHODS AND MATERIALS

20 **Immunization Protocols and Splenocyte /PBMC Preparation**

Immunizations: Mice were immunized with the selected adenoviruses diluted in 0.1 ml of buffer. Each vector dose was divided in two aliquot of 50 µl and injected in both quadriceps of mice.

25 Splenocyte Preparation: Mice were sacrificed 3 weeks post-injection and their spleens excised and transferred in 10 ml of R10 (10% FCS, 55mM 2-mercaptoethanol, 1M HEPES buffer, 2mM L-glutamine, 1X penicillin-streptomycin solution in RPMI medium 1640). Spleens were minced through a steel screen and, after the screen was washed with 2 ml of R10, splenocytes were transferred in a 50 ml Falcon tube and centrifuged at 1200 rpm, 10 min, room
30 temperature (rt). Supernatant was removed and 3 ml of ACK lysis buffer (Gibco BRL Formulation#79-0422DG) were added. Cells were incubated 5 min, rt. 45 ml of 1X PBS were

added and tubes were centrifuged as above. After washing with 30 ml of R10, cells were resuspended in 5 ml of R10, filtered through a 70 m Nylon cell strainer (Falcon 2350). 10 μ l of cells were diluted with 990 μ l Turk's solution (Merck 040417345) and counted. Cells were finally diluted to 10⁷ cells/ml in R10.

- 5 Peripheral blood mononuclear cell (PBMC) preparation: Mice blood samples (150 μ l) were transferred to 2ml eppendorf tubes with 50 μ l PBS/2% EDTA. 1 ml ACK buffer was added to each tube. Gently mixed and incubated at RT for 5 min. Samples were centrifuged at 1500rpm in microcentrifuge for 5 min. Supernatant was discharged white cell pellets deriving from the same immunized cohorts were combined. ACK buffer incubation was repeated then
10 pellets of PBMC were resuspended in 1 ml of R10 medium.

IFN- γ ELISPOT Assay

- Millipore MAIP 45 plates were coated with 100 μ l/well of purified rat anti-mouse IFN- γ monoclonal antibody (Pharmingen, cat. 551216) diluted at 2.5 μ g/ml in PBS and incubated
15 over-night (o/n) at 4⁰C. Plates were washed 2X with sterile PBS and un-specific binding sites were blocked by incubation for 2hrs in the CO₂ incubator with 200 μ l/well of R10. In the immunization experiments with Ad vectors expressing HIV gag, a 9-mer peptide (AMQMLKETI, a CD8 HIV gag epitope mapped in Balb/C mice) (SEQ ID NO: 47) was diluted to 2 μ g/ml in R10 and added to the wells in the amount of 50 μ l/well. In immunization
20 experiments conducted with HCV-NS expressing vectors, a pool of peptides covering NS3 helicase domain as well a 9-mer peptide representing a mapped CD8 epitope comprised in helicase domain were used. Immunization experiments with ChAds expressing human CEA antigen were evaluated by pools of overlapping 15-mer peptides covering the entire amino acid sequence. As controls DMSO and Concanavalin A were used. Cells were added to each well at
25 the amount of 5X10⁵ and 2.5X10⁵. After an o/n incubation in the CO₂ incubator, plates were washed with 0.05% Tween 20/PBS and 50 μ l/ well of biotinylated rat anti-mouse IFN- γ monoclonal antibody (PharMingen cat. 554410) diluted 1/250 in assay buffer (5%FBS, 0.005% Tween20, PBS) were added. Plates were incubated o/n at 4⁰C and washed as above. Streptavidin-alkaline phosphatase conjugate (BD554065) was diluted 1 /2500 in assay buffer and
30 added in the amount of 50 μ l/well for 2 hrs rt. After washing, plates were developed adding 50

μl/well of BCIP/NBT1-step solution (Pierce 34042). Reaction was stopped by washing wells with deionized water. Spots were automatically counted by an ELISPOT reader.

Murine IFN-γ Intracellular Staining (ICS)

5 Splenocytes were diluted at 2×10^6 cells in 1 ml of R10 and stimulated with the same antigens described above at the concentration of 2 μg/ml. As controls, DMSO and Staphylococcal Enterotoxin B (SEB) were used. After an overnight incubation in the CO₂ incubator, cells were washed with FACS buffer (1% FCS, 0.01% NaN₃, PBS) and purified anti-mouse CD16/CD32 Fc block (clone 2.4G2, Pharmingen cat. 553142) was diluted 1/25, added in
10 the amount of 100 μl/sample and incubated for 15 min at 4°C. Cells were washed in FACS buffer and APC conjugated anti-mouse CD3e (clone 145-2C11, Pharmingen #553066), PE conjugated anti-mouse CD4 (clone L3T4, BD Pharmingen cat. 553142) and PerCP conjugated anti-mouse CD8a (clone 53-6.7, Pharmingen cat. 553036) diluted 1:50 in FACS buffer were added in the amount of 100 μl/sample. Cells were incubated 30 min rt, washed, fixed and permeabilized
15 (Becton Dickinson, FACS Perm 2) and incubated with FITC conjugated anti-mouse IFN-γ Pharmingen cat. 554411) diluted 1:50 in PermWash (100 ul/sample) for 30 min at RT. After washing cells were resuspended in 500 ul 1% formaldehyde/PBS and intracellular cytokine staining (ICS) analyzed on a FACS-Calibur flow cytometer, using CellQuest software (Becton Dickinson).

20

EXAMPLE 4 ChAd VECTORS ELICIT STRONG CMI RESPONSES IN MICE

The ability of the ChAd vectors disclosed herein to elicit a cell-mediated immune response (CMI) was evaluated in mice using vectors expressing an HIV gag transgene. Briefly, groups of 5 Balb/C mice were injected with ten-fold increasing doses of the different vectors
25 starting from 10^5 up to 10^{10} vp/mouse.

The strength of the immune response was determined three weeks after the injection by quantifying gag-specific CD8+ T cells in the splenocytes. The number of IFN-γ secreting CD8+ T cells was determined by ELISPOT assay or by IFN-γ intracellular staining and FACS analysis after stimulation *in vitro* with a peptide reproducing a gag CD8+ T cell
30 epitope mapped in Balb/C mice.

The results obtained from the 5 immunized animals, reported in Table 3, are expressed as spot forming cells per 10^6 splenocytes. Shown are the number of spot forming cells per million splenocytes following incubation with 9-mer CD8+ gag epitope or with gag peptide pool. The gag peptide pool consisted of 20-aa peptide overlapping by 10aa encompassing the
35 entire gag sequence. Positive values are reported in bold.

5 The data provided in Table 3 indicate that the administration of the ChAd vectors disclosed and claimed herein elicits a strong cell mediated immune response which is comparable to the response elicited by hAd5. By looking at the lowest vector dose resulting in a positive immunization result (immunization breakpoint), we ranked the potency of the different vectors being subgroup C ChAd3gag the most potent with a breakpoint at 10^6 pp vector dose. Ranking by immunization break-points is shown in Figure 33.

Table 3. Gag-specific T cell response in BalbC mice immunized with chimpanzee Ad vectors.

Vaccination	10 ⁴ vp		10 ⁶ vp		10 ⁷ vp		10 ⁸ vp		10 ⁹ vp		10 ¹⁰ vp	
	mock	Gag	mock	Gag	mock	Gag	mock	Gag	mock	Gag	mock	Gag
ChAd3DE1gag	1	neg	1	944	1	1298	1	1258	NT	NT	NT	NT
	3	neg	1	1038	1	1858	1	1962	NT	NT	NT	NT
	1	neg	1	859	1	1823	1	1931	NT	NT	NT	NT
	1	neg	1	1820	1	1388	1	1389	NT	NT	NT	NT
	1	neg	1	1820	5	1442	4	1438	NT	NT	NT	NT
CV32DE1gag	NT	NT	1	neg	2	475	1	2910	NT	NT	NT	NT
	NT	NT	1	neg	1	433	1	401	NT	NT	NT	NT
	NT	NT	1	neg	1	243	1	634	NT	NT	NT	NT
	NT	NT	1	neg	1	505	2	3457	NT	NT	NT	NT
	NT	NT	1	neg	1	553	2	1684	NT	NT	NT	NT
CV88DE1gag	NT	NT	3	neg	1	340	2	332	0	408	2	635
	NT	NT	1	neg	1	512	0	538	1	258	3	1172
	NT	NT	0	neg	2	458	3	944	2	462	2	505
	NT	NT	7	neg	0	148	1	519	0	488	2	1184
	NT	NT	0	neg	2	1418	1	243	0	240	1	789
ChAd9DE1gag	NT	NT	1	neg	7	369	1	609	NT	NT	NT	NT
	NT	NT	1	neg	1	508	1	739	NT	NT	NT	NT
	NT	NT	1	neg	1	299	18	291	NT	NT	NT	NT
	NT	NT	1	neg	2	507	8	926	NT	NT	NT	NT
	NT	NT	0.5	neg	1	36	40	1034	NT	NT	NT	NT
ChAd10DE1gag	NT	NT	1	neg	1	83	1	622.5	NT	NT	NT	NT
	NT	NT	1	neg	1	42.5	1	1033	NT	NT	NT	NT
	NT	NT	1	neg	1	48	1	1339.5	NT	NT	NT	NT
	NT	NT	1	neg	4	51	1	1132	NT	NT	NT	NT
	NT	NT	1	neg	1	466.5	1	521.5	NT	NT	NT	NT
ChAd6DE1gag	NT	NT	1	neg	1	34	1	721	NT	NT	NT	NT
	NT	NT	1	neg	10	4	1	560	NT	NT	NT	NT
	NT	NT	1	neg	1	24	1	624	NT	NT	NT	NT
	NT	NT	1	neg	1	225	3	3002	NT	NT	NT	NT
	NT	NT	1	neg	1	276	4	1738	NT	NT	NT	NT
ChAd11DE1gag	1	neg	1	neg	0	573	NT	NT	NT	NT	NT	NT
	0	neg	0	neg	0	919	NT	NT	NT	NT	NT	NT
	0	neg	1	neg	1	1438	NT	NT	NT	NT	NT	NT
	2	neg	0	neg	0	0	NT	NT	NT	NT	NT	NT
	1	neg	1	neg	0	456	NT	NT	NT	NT	NT	NT
ChAd20DE1gag	0	neg	0	neg	0	1	NT	NT	NT	NT	NT	NT
	2	neg	0	neg	0	408	NT	NT	NT	NT	NT	NT
	0	neg	0	neg	0	414	NT	NT	NT	NT	NT	NT
	1	neg	0	neg	0	2	NT	NT	NT	NT	NT	NT
	0	neg	0	neg	1	311	NT	NT	NT	NT	NT	NT
ChAd7DE1gag	NT	NT	1	neg	1	neg	1	1044	NT	NT	NT	NT
	NT	NT	3	neg	1	neg	1	806	NT	NT	NT	NT
	NT	NT	1	neg	8	neg	1	407	NT	NT	NT	NT
	NT	NT	1	neg	1	neg	2	567	NT	NT	NT	NT
	NT	NT	1	neg	3	neg	1	1677	NT	NT	NT	NT
CV32DE1gag	NT	NT	NT	NT	1	neg	0	83	0	291	0	194
	NT	NT	NT	NT	3	neg	0	382	0	805	2	380
	NT	NT	NT	NT	0	neg	1	97	0	138	1	501
	NT	NT	NT	NT	1	neg	5	96	4	1162	0	1115
	NT	NT	NT	NT	2	neg	1	328	NT	NT	0	596
ChAd4DE1gag	NT	NT	1	neg	0	neg	0	0	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	0	159	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	0	1	NT	NT	NT	NT
	NT	NT	1	neg	0	neg	0	234	NT	NT	NT	NT
	NT	NT	1	neg	0	neg	1	0	NT	NT	NT	NT
ChAd16DE1gag	NT	NT	0	neg	0	neg	0	243	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	1	296	NT	NT	NT	NT
	NT	NT	0	neg	2	neg	1	88	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	1	433	NT	NT	NT	NT
	NT	NT	1	neg	0	neg	1	28	NT	NT	NT	NT

EXAMPLE 5 ChAd3 AND CV33 GAG VECTORS ELICIT A CMI RESPONSE CHARACTERIZED BY GAG-SPECIFIC CD8+ T CELLS

In order to characterize the CMI response elicited in response to the ChAd vectors comprising HIV gag transgene, splenocytes pooled from cohorts of five mice immunized with different doses of vector were analyzed by intracellular IFN- γ staining. The data shown in table 3 and table 4 were collected in separate experiments.

Splenocytes were diluted at 2×10^6 cells in 1 ml of R10 and stimulated with the same antigens described above at the concentration of 2 $\mu\text{g/ml}$. As controls, DMSO and SEB (Staphylococcal Enterotoxin B) were used. After an o/n incubation in the CO_2 incubator, cells were washed with FACS buffer (1% FCS, 0.01% NaN_3 , PBS) and purified anti-mouse CD16/CD32 Fc block (clone 2.4G2, Pharmingen cat. 553142) was diluted 1/25, added in the amount of 100 $\mu\text{l/sample}$ and incubated for 15min at 4°C . Cells were washed in FACS buffer and APC conjugated anti-mouse CD3e (clone 145-2C11, Pharmingen #553066), PE conjugated anti-mouse CD4 (clone L3T4, BD Pharmingen cat. 553142) and PerCP conjugated anti-mouse CD8a (clone 53-6.7, Pharmingen cat. 553036) diluted 1:50 in FACS buffer were added in the amount of 100 $\mu\text{l/sample}$. Cells were incubated 30 min rt, washed, fixed and permeabilized (Becton Dickinson, FACS Perm 2) and incubated with FITC conjugated anti-mouse IFN- γ Pharmingen cat. 554411) diluted 1:50 in PermWash (100 $\mu\text{l/sample}$) for 30 min at RT. After washing cells were resuspended in 500 μl 1% formaldehyde/PBS and analyzed on a FACS-Calibur flow cytometer, using CellQuest software (Becton Dickinson).

Table 4 provides data summarizing the percentage of gag-specific CD3+T cells that were either gag-specific CD8+ or CD4+ T cells. Positive results are reported in bold. The data provided herein indicate that the cellular profile of the immune response elicited by ChAd vectors derived from viruses classified into different serotype subgroups (i.e., subgroups C, D and E) are similar and all of the gag-specific responses characterized predominantly by CD8+ T cells. In addition, it is noted that at high vector doses a gag-specific CD4+ response becomes evident in all immunization experiments. The ICS assay confirmed that ChAd3 vector can stimulate anti-gag CD8+ response at 10^6 vector dose.

Table 4. Characterization of gag-specific T cells in mice immunized with Chimp adenovirus vectors of different subgroups.

vaccine		10 ⁵		10 ⁶		10 ⁷		10 ⁸		10 ⁹	
		DMSO	gag	DMSO	gag	DMSO	gag	DMSO	gag	DMSO	gag
ChAd3DE1gag	%CD8 ⁺ CD3 ⁺	NT	NT	0.01%	4.65%	0.01%	17.15%	0.04%	24.71%	NT	NT
	%CD4 ⁺ CD3 ⁺	NT	NT	0.00%	0.07%	0.03%	0.08%	0.04%	0.28%	NT	NT
CV33DE1gag	%CD8 ⁺ CD3 ⁺	NT	NT	0.02%	0.01%	0.01%	0.83%	0.03%	8.69%	NT	NT
	%CD4 ⁺ CD3 ⁺	NT	NT	0.00%	0.00%	0.00%	0.04%	0.01%	0.10%	NT	NT
ChAd9DE1gag	%CD8 ⁺ CD3 ⁺	NT	NT	0.02%	0.01%	0.01%	0.68%	NT	NT	0.04%	4.73%
	%CD4 ⁺ CD3 ⁺	NT	NT	0.00%	0.00%	0.00%	0.00%	NT	NT	0.00%	0.01%
ChAd10DE1gag	%CD8 ⁺ CD3 ⁺	NT	NT	0.02%	0.01%	0.01%	0.57%	NT	NT	0.02%	5.04%
	%CD4 ⁺ CD3 ⁺	NT	NT	0.00%	0.00%	0.00%	0.00%	NT	NT	0.00%	0.01%
ChAd6DE1gag	%CD8 ⁺ CD3 ⁺	NT	NT	0.00%	0.01%	0.00%	0.59%	0.01%	14.28%	NT	NT
	%CD4 ⁺ CD3 ⁺	NT	NT	0.00%	0.00%	0.00%	0.05%	0.01%	0.12%	NT	NT
ChAd7DE1gag	%CD8 ⁺ CD3 ⁺	NT	NT	0.01%	0.02%	0.01%	0.00%	0.02%	5.00%	NT	NT
	%CD4 ⁺ CD3 ⁺	NT	NT	0.00%	0.01%	0.00%	0.00%	0.01%	0.21%	NT	NT

EXAMPE 6 ChAd VECTORS ELICIT HCV NS-SPECIFIC T CELL RESPONSE

The potency of CV32-NSmut and CV33-NSmut vectors was evaluated in C57/Black6 mice relative to the potency of MRKAd6NSmut. The animals were injected with 10-fold increasing doses of vector starting from 10⁷ up to 10⁹ vp/mouse. CMI was analyzed 3 weeks after a single injection by IFN-γ ELISPOT and IFN-γ intracellular staining by stimulating T cells with a 9-mer peptide reproducing a CD8⁺ T cell epitope mapped in the helicase domain of NS3 protein. The data provided in Table 5 summarize the number of spot-forming cells per million splenocytes following incubation in absence (mock) or in presence of NS3 9-mer peptide.

The data indicate that both CV32 and CV33 vectors expressing HCV-NS stimulate strong T cell responses. Based on the observation that the first positive result for the CV32 vector was obtained by injecting 10⁹ vp/dose, the immunization potency of CV32DE1E3 NSmut vector appears to be approximately 100-fold lower than human subgroup C Ad6DE1E3 NSmut vector. The parallel experiment with MRKAd6NSmut indicated that a dose of 10⁷ vp/animal was sufficient to stimulate cell mediated immunity. Therefore, these results confirm the lower immunization potency of CV32-derived vectors relative to human subgroup C vectors

(such as hAd5 and hAd6) that was also observed in the experiment with gag expressing vectors (see Table 3).

Table 5. HCV NS-specific T cell response in mice immunized with MRKAd6 NSmut, CV32NSmut or CV33NSmut

Vaccination	10 ⁷ vp		10 ⁸ vp		10 ⁹ vp		10 ¹⁰ vp	
	Mock	NS3	mock	NS3	mock	NS3	mock	NS3
MRKAd6NSmut	1	345	1	449	NT	NT	NT	NT
	1	248	1	1590	NT	NT	NT	NT
	1	1	1	549	NT	NT	NT	NT
	1	262			NT	NT	NT	NT
					NT	NT	NT	NT
CV33NSmut	1	1	1	195	2	338	NT	NT
	1	2	1	409	1	1136	NT	NT
	1	1	1	396	1	497	NT	NT
	1	2	2	172	1	344	NT	NT
	1	237			1	163	NT	NT
CV32NSmut	neg	neg	1	181	1	118	1	176
	neg	neg	1	71	1	239	1	238
	neg	neg	1	56	1	862	1	555
	neg	neg	1	459	1	219	1	545
	neg	neg	1	195	1	123	1	578

EXAMPLE 7 ANTI-Ad5 PRE-EXISTING IMMUNITY DOES NOT ABROGATE ANTI-GAG CMI ELICITED BY ChAd3gag

To evaluate the impact on ChAd3 immunization of the pre-existing immunity against the high seroprevalent Ad5, 4 cohorts of 5 BalbC mice were pre-immunized with two injection of 10¹⁰ vp of Ad5 wt in the quadriceps at week 0 and 2. As control, 2 cohorts of 5 mice were injected at the same time points with buffer only. Cohorts of Ad5 pre-immunized mice were then immunized with 10⁶ and 10⁷ vp/mouse of either Ad5gag or ChAd3gag vectors. Cohorts of control (naïve) mice were immunized with 10⁶ vp/mouse of Ad5gag or ChAd3gag vectors. Anti-Ad5 and ChAd3 neutralizing immunity was evaluated at week 4 by the neutralization assay described above using Ad5 and ChAd3 SEAP vectors. Anti-gag immunity was evaluated by ELISPOT analysis on purified splenocytes stimulated with gag 9-mer peptide containing a gag epitope mapped in BalbC mice. The results reported in figure 36 demonstrated that Anti-Ad5 immunity does not abrogate anti-gag CMI elicited by ChAd3gag while, as expected, anti-Ad5 immunity completely block Ad5gag immunization.

EXAMPLE 8 ChAd3hCEA IMMUNIZATION ELICITS A STRONG CEA-SPECIFIC IMMUNE RESPONSE IN TRANSGENIC MICE EXPRESSING HUMAN CEA

5 The ability of the ChAd vectors disclosed and claimed herein to elicit an immune response against a self-antigen therefore breaking the tolerance was also evaluated in transgenic mice expressing human CEA (Clarke, P *et al. Cancer Res.* (1998) 58(7):1469-77.)

10 Cohorts of 8 mice were injected in the quadriceps with 10^{10} vp of ChAd3hCEA or Ad5hCEA as already described. The immune response against CEA was followed weekly up to day 75 on PBMC stimulated with a pool of 15-mer peptides encompassing human CEA aminoacid sequence from aa 497 to the end (aa 703). Anti-CEA immunity was evaluated by ICS determining CD4-CD8+ T cells secreting interferon- γ in response to CEA peptide pool incubation.

15 The results reported in figure 37 demonstrate that ChAd3hCEA vector immunization stimulate a more sustained CD8+ T cell response against human CEA than Ad5 expressing the same transgene.

PRIMATE IMMUNIZATION STUDIES

METHODS AND MATERIALS

Immunization Protocol

20 The ability of the ChAd vectors disclosed and claimed herein to elicit CMI in Rhesus macaques (referred to herein as monkeys) was also evaluated. The macaques were anesthetized (ketamine/xylazine) and the vaccines were delivered i.m. in 0.5-mL aliquots into both deltoid muscles using tuberculin syringes (Becton-Dickinson). In all cases the macaques were between 3-10 kg in weight, and the total dose of each vaccine was administered in 1 mL of
25 buffer.

Sera and peripheral blood mononuclear cells (PBMC) were prepared from blood samples collected at several time points during the immunization regimen. All animal care and treatment were in accordance with standards approved by the Institutional Animal Care and Use Committee according to the principles set forth in the *Guide for Care and Use of Laboratory*
30 *Animals*, Institute of Laboratory Animal Resources, National Research Council.

ELISPOT Assay

The IFN- γ ELISPOT assays for rhesus macaques were conducted following a previously described protocol (Allen *et al.*, 2001 *J. Virol.* 75(2):738-749), with some modifications. For gag-specific stimulation, a peptide pool was prepared from 20-aa peptides that encompass the entire HIV-1 gag sequence with 10-aa overlaps (Synpep Corp., Dublin, CA). For HCV NS-specific stimulation 6 peptide pools were prepared from 15-aa peptides that encompass the entire HCV-NS sequence from NS3 to NS5b with 10-aa overlaps. HER2/neu and CEA-specific stimulations were performed with 15-aa peptides that encompass the entire protein sequence with 10-aa overlaps.

To each well, 50 μ L of $2-4 \times 10^5$ peripheral blood mononuclear cells (PBMCs) were added; the cells were counted using Beckman Coulter Z2 particle analyzer with a lower size cut-off set at 80 fL. Either 50 μ L of media or the gag peptide pool at 8 μ g/mL concentration per peptide was added to the PBMC. The samples were incubated at 37°C, 5% CO₂ for 20-24 hrs. Spots were developed accordingly and the plates were processed using custom-built imager and automatic counting subroutine based on the ImagePro platform (Silver Spring, MD); the counts were normalized to 10^6 cell input.

Intracellular Cytokine Staining (ICS)

To 1 ml of 2×10^6 PBMC/mL in complete RPMI media (in 17x100mm round bottom polypropylene tubes (Sarstedt, Newton, NC)), anti-hCD28 (clone L293, Becton-Dickinson) and anti-hCD49d (clone L25, Becton-Dickinson) monoclonal antibodies were added to a final concentration of 1 μ g/mL. For gag-specific stimulation, 10 μ L of the peptide pool (at 0.4 mg/mL per peptide) were added. Similar conditions were used for HCV NS-specific stimulation. The tubes were incubated at 37 °C for 1 hr., after which 20 μ L of 5 mg/mL of brefeldin A (Sigma) were added. The cells were incubated for 16 hr at 37 °C, 5% CO₂, 90% humidity. 4 mL cold PBS/2%FBS were added to each tube and the cells were pelleted for 10 min at 1200 rpm. The cells were re-suspended in PBS/2%FBS and stained (30 min, 4 °C) for surface markers using several fluorescent-tagged mAbs: 20 μ L per tube anti-hCD3-APC, clone FN-18 (Biosource); 20 μ L anti-hCD8-PerCP, clone SK1 (Becton Dickinson, Franklin Lakes, NJ); and 20 μ L anti-hCD4-PE, clone SK3 (Becton Dickinson). Sample handling from this stage was conducted in the dark. The cells were washed and incubated in 750 μ L 1xFACS Perm buffer (Becton Dickinson) for 10 min at room temperature. The cells were pelleted and re-suspended in PBS/2%FBS and 0.1 μ g of FITC-anti-hIFN- γ , clone MD-1 (Biosource) was added. After 30 min incubation, the cells were washed and re-suspended in PBS. Samples were analyzed using all four color channels of the Becton Dickinson FACSCalibur instrument. To

analyze the data, the low side- and forward-scatter lymphocyte population was initially gated; a common fluorescence cut-off for cytokine-positive events was used for both CD4⁺ and CD8⁺ populations, and for both mock and gag-peptide reaction tubes of a sample.

5 EXAMPLE 9 A HOMOLOGOUS PRIME-BOOST REGIMEN USING ChAd ΔE1-gag VECTORS ELICITS GAG-SPECIFIC T CELLS IN MONKEYS

Cohorts of 3 animals were given intramuscular injection at week 0 and week 4 of either of the following constructs: 10¹⁰ vp of CV-32ΔE1-gag; or 10¹⁰ vp CV33ΔE1-gag; or 10¹⁰ vp and 10⁸ vp MRKAd5ΔE1gag. PBMCs collected at regular 4-wks intervals were analyzed in an ELISPOT assay. The results provided in Table 6, which indicate the number of spot-forming cells per million PBMC following incubation in absence (mock) or presence of Gag peptide pool establish that both CV32ΔE1-gag and CV-33ΔE1gag are able to induce significant levels of gag-specific T cells in non-human primates. It is interesting to note that after a single dose (wk 4), the CV32ΔE1-gag responses were comparable to MRKAd5 ΔE1-gag 10⁸ vp dose and lower than that of MRKAd5-gag 10¹⁰ vp/dose. CV33ΔE1-gag 10¹⁰ vp/dose induces a response comparable to that of MRKAd5-gag 10¹⁰ vp/dose. This result was confirmed at week 8 after the second dose.

Table 6. Gag-specific T cell response in monkeys immunized with MRKAd5 ΔE1-gag, CV32ΔE1-gag, CV33ΔE1-gag.

Vaccination T=0	vector dose	Monk #	Pre-bleed		T=4		T=8	
			Mock	Gag	Mock	Gag	Mock	Gag
CV32ΔE1gag	10 ¹⁰ vp	01C023	1	0	14	353	3	278
		01C029	1	3	13	605	3	419
		01C032	1	0	5	274	1	179
CV33ΔE1gag	10 ¹⁰ vp	01C033	0	0	9	1545	1	659
		01C036	4	5	4	1540	13	881
		01D303	0	3	19	949	10	628
MRKAd5gag	10 ⁸ vp	01D267	0	0	4	473	0	341
		01D279	1	4	44	831	6	336
		01D284	4	5	4	264	5	129
MRKAd5gag	10 ¹⁰ vp	99C218	0	3	5	2500	0	1580
		99C227	6	1	4	529	5	365
		99D185	ND	ND	0	425	0	310

EXAMPLE 10 ChAd VECTORS ELICIT A HCV NS-SPECIFIC T-CELL RESPONSE IN A HETEROLOGOUS PRIME-BOOST REGIMEN

In a separate experiment, groups of two and three monkeys were given immunization at week 0, 4 of MRK Ad6NSoptmut vector at 10^8 or 10^{10} vp per animal. The animals were boosted with the same virus at the same dose at week 24 and then boosted again at week 104 with CV33-NSmut at 10^{10} vp per animal. The results are presented in Tables 7 and 8 which summarize the number of spot-forming cells per million PBMC following incubation in absence (mock) or presence of HCV NS peptide pool.

T cell immunity, as assessed by IFN- γ ELISPOT, showed a peak response at week 4 after the first dose in the animals injected with 10^{10} vp (Table 8) and at week 8 (post-dose 2) in the animals injected at 10^8 (Table 7). The response was not boosted by the injection at week 24 ("homologous boost"), while a strong boost effect was observed after the injection with CV33-NSmut ("heterologous boost").

Table 7. HCV NS-specific T cell response in monkeys immunized with MRK Ad6NSoptmut at 10^8 vp/animal and boosted with CV33-NSmut.

Vaccine	MRKAd6NSoptmut 10^8 vp								CV33-NSmut 10^{10} vp			
	post-priming I dose		post-priming II dose		pre-homologous boost		post-homologous boost		pre-heterologous boost		post-heterologous boost	
	T=4		T=8		T=24		T=28		T=104		T=108	
monkey	95116	138T	95116	138T	95116	138T	95116	138T	95116	138T	95116	138T
poolF	44	112	77	124	115	176	105	55	120	150	188	2228
poolG	20	2110	88	1975	201	1105	94	884	120	192	96	4590
poolH	12	18	54	22	169	221	28	9	81	33	447	543
poolI	14	53	62	47	163	189	98	18	80	67	71	515
poolL	33	88	58	44	353	608	235	33	110	131	224	308
poolM	184	75	168	138	204	336	67	44	55	46	2028	1570
DMSO	14	3	44	7	104	79	33	6	57	40	33	65

Table 8. HCV NS-specific T cell response in monkeys immunized MRK Ad6NSoptmut at and 10^{10} vp/animal and boosted with CV33-NSmut.

Vaccine	MRKAd6NSoptmut 10^{10} vp												CV33-NSmut 10^{10}					
	post-priming I dose T=4			post-priming II dose T=8			pre-homologous boost T=24			post-homologous boost T=28			pre-heterologous boost T=104			post-heterologous boost T=108		
time point	98D209	106Q	113Q	98D209	106Q	113Q	98D209	106Q	113Q	98D209	106Q	113Q	98D209	106Q	113Q	98D209	106Q	113Q
monkey																		
poolF	3110	263	404	1340	300	723	678	61	583	321	123	1438	204	192	328	1581	1525	1714
poolG	2115	642	1008	1070	316	2205	685	71	701	251	178	1758	166	106	625	1118	524	4238
poolH	373	72	19	358	43	43	424	24	42	51	23	18	92	45	55	413	58	211
poolI	103	37	347	80	36	531	237	39	169	12	35	485	66	79	376	459	85	2738
poolL	149	22	10	93	36	29	279	46	48	11	49	51	89	109	73	199	76	431
poolM	314	428	19	153	243	20	333	81	38	38	134	11	41	81	9	228	1440	227
DMSO	0	1	3	18	16	5	128	8	9	8	10	16	20	51	12	18	13	5

The efficiency of heterologous boost with chimp Ad vectors was evaluated in a second experiment. Cohorts of three monkeys were immunized at week 0 and week 4 with MRKAd5gag (10^{10} vp/animal), MRKAd6NSmut (10^{10} vp/animal) or with the combination of both vectors (10^{10} vp/animal each vector) then boosted with the same immunogen at week 24 (homologous boost). Homologous boost was performed with the same immunogens; heterologous boost was performed with CV33gag, CV32 NSmut or with the two vectors in combination. The results provided in Table 9 summarize the number of spot-forming cells per million PBMC following incubation in absence (mock) or presence of HCV NS peptide pool.

The same cohorts were boosted again at week 51 with CV33gag (10^{10} vp/animal), CV32NSmut (10^{10} vp/animal) and with the combination of the two vectors (10^{10} vp/animal each vector). The results provided in Table 9 further indicate that the homologous boost was not efficient since the responses are below the peak observed at week 4 after the injection of the first dose of vaccine. A strong boosting effect was measured by IFN- γ ELISPOT at week 54 after immunization with heterologous chimp vectors.

Table 9. Immunization with Chimp Ad vectors efficiently boost Gag and HCV NS-specific T cell response in monkeys immunized with MRK Ad5gag or MRK Ad6NSoptmut at 10^4 vp/animal.

Vaccine	MRKAd5gag									CV33gag								
	post-dose 1 (T=4)			post-dose 2 (T=8)			pre-homol. boost (T=24)			post-homol. boost (T=28)			pre-heterol. boost (T=51)			post-heterol. boost (T=54)		
animal ID	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299
poolF	18	35	60	16	29	14	37	76	40	37	8	14	37	27	44	43	44	70
poolG	16	23	49	4	28	31	54	95	106	81	2	46	36	27	37	84	108	109
poolH	45	51	57	18	31	42	55	88	55	47	11	32	69	36	60	85	58	120
poolI	21	21	48	4	26	11	19	54	26	38	6	6	22	11	32	33	26	24
poolL	15	21	58	9	31	20	71	183	128	106	6	27	61	21	65	28	45	44
poolM	39	24	48	26	14	49	38	93	39	59	6	19	62	23	38	27	19	14
Gag	1764	2208	2762	574	1906	1959	391	935	702	2123	336	736	485	833	1384	4003	4333	3863
DMSO	9	13	37	7	14	13	16	76	33	26	3	11	28	19	39	23	16	53

Vaccine	MRK Ad5gag + MRKAd6NSmut									CV33gag + CV32NSmut								
	post-dose 1 (T=4)			post-dose 2 (T=8)			pre-homol. boost (T=24)			post-homol. boost (T=28)			pre-heterol. boost (T=51)			post-heterol. boost (T=54)		
animal ID	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240
poolF	438	118	105	720	116	154	206	108	242	408	99	219	778	135	56	1701	1121	424
poolG	21	784	1483	44	362	940	19	234	548	47	781	844	78	363	285	228	3180	2770
poolH	24	53	8	46	27	19	13	66	93	49	41	87	115	50	28	97	291	104
poolI	83	28	9	90	24	8	16	40	68	33	16	42	56	19	8	165	145	22
poolL	13	14	13	16	17	9	28	101	140	39	27	78	59	28	15	137	815	463
poolM	39	31	6	101	27	18	21	73	107	44	26	78	114	28	10	219	109	21
Gag	2138	1044	1063	2260	505	819	454	241	456	1100	368	716	1542	237	161	4460	2908	1764
DMSO	5	6	3	8	5	1	10	18	43	9	13	28	14	18	12	9	21	6

Vaccine	MRKAd6 NSmut									CV32NSmut								
	post-dose 1 (T=4)			post-dose 2 (T=8)			pre-homol. boost (T=24)			post-homol. boost (T=28)			pre-heterol. boost (T=51)			post-heterol. boost (T=54)		
animal ID	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159
poolF	139	44	82	92	121	63	62	116	54	44	42	23	57	85	53	313	385	261
poolG	154	253	119	77	156	108	93	165	126	104	59	39	44	198	48	196	764	559
poolH	1284	41	211	768	35	124	394	84	77	24	817	48	624	31	116	3758	90	925
poolI	302	22	1174	221	16	1069	134	31	561	18	133	478	84	16	362	485	51	2951
poolL	28	16	48	35	32	21	141	113	78	19	48	17	46	33	46	379	339	541
poolM	1329	1007	36	579	392	30	314	293	43	558	398	22	159	369	33	1278	1750	16
Gag	15	9	7	13	5	2	36	33	36	9	23	14	16	8	10	37	9	26
DMSO	16	4	5	9	6	4	23	17	8	1	9	3	23	8	6	26	9	10

EXAMPLE 11 VACCINATION WITH A ChAd VECTOR COMPRISING A TAA BREAKS TOLERANCE AND ELICITS A TAA-SPECIFIC T CELL RESPONSE IN MONKEYS

Experiments designed to determine whether chimpanzee adenoviral vectors are sufficiently immunogenic to break the tolerance to a self-antigen and to document the utility of chimpanzee vectors for boosting an immune response primed with a human adenoviral vector were performed in cohorts of four monkeys. Animals were immunized with three injection at

week 0, 2 and 4 of Ad5DE1 RhCEA (10^{11} vp), comprising the tumor associated antigen CEA, followed by vaccination at week 16, 18 and 20 with CV33DE1 RhCEA (10^{11} vp). T cell response was measured by IFN γ ELISPOT with rhesus CEA peptides.

5 The results reported in figure 34, which provide the number of spot-forming cells per million PBMC following incubation in absence (DMSO) or in presence of rhesus CEA C and D peptides pools, establish that an immunization protocol based on vaccination with two different Ad serotypes leads to a sustained T cell response against CEA in non-human primates.

10 While the invention has been described in detail with reference to certain preferred embodiments thereof, it will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

CLAIMS

What is claimed is:

- 5 1. An isolated chimpanzee nucleic acid sequence selected from the group consisting of:
 - a) SEQ ID NO:1
 - b) SEQ ID NO: 2; and
 - c) a nucleic acid sequence complementary to the sequence of (a) or (b).
- 10 2. An isolated recombinant chimpanzee serotype comprising any combination of hexon and fiber nucleic acid sequences selected from the groups of:
 - a) hexon sequences SEQ ID NOS: 16- 25; and
 - b) fiber sequences SEQ ID NOS: 6-15. ,
- 15 3. A replication defective chimpanzee adenoviral (ChAd) vector comprising the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:2 and a transgene which encodes at least one immunogen operatively linked to regulatory sequences which direct expression of said transgene in mammalian cells, wherein said vector lacks the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:2 which comprises at least one gene selected from the group consisting of
20 adenoviral E1, E2, E3, and E4.
4. A replication defective ChAd vector which comprises a deletion/disruption in the E1 nucleotide sequence in the region from bp 460 to bp 3542 of SEQ ID NO: 1 or from bp 457 to bp 3425 of SEQ ID NO:2.
25
5. The ChAd vector according to claim 4, wherein the vector comprises a transgene selected from the group consisting of: HIV, HBV, HCV, HPV, HSV1, HSV2, SARS CoV, *Plasmodium malariae*, Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, and *Mycobacterium tuberculosis*.
30
6. The ChAd vector according to claim 4, wherein the vector comprises a deletion/disruption in the E1 nucleotide sequence in the region from bp 460 to bp 3542 of SEQ ID NO: 1 or from bp 457 to bp 3425 of SEQ ID NO:2 and further wherein the vector comprises a transgene encoding at least one tumor associated antigen (TAA).
35

7. The ChAd vector according to claim 6 wherein the at least one TAA is selected from the group consisting of: HER2 NEU, CEA, EPCAM, PSA, PSMA, TELOMERASE, GP100, MELAN-A/MART-1, MUC-1, NY-ESO-1, SURVIVIN, STROMELYSIN 3, TYROSINASE, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, HKLP2, 5T4 and VEGFR2.

8. A host cell comprising a nucleic acid molecule according to claim 1 or claim 2, wherein said host cell expresses one or more adenoviral regions selected from the group consisting of E1a, E1b, E2a, E2b, E4 orfs 1, 2, 3, 4, 5, 6, 6/7, pIX, IVa2, regions L1, L2, L3, L4, L5.

9. A method of producing a replication-defective chimpanzee adenoviral vector comprising introducing an adenoviral vector according to Claims 3 into an adenoviral E-1 expressing human cell, and harvesting the resulting adenoviruses.

10. The method according to Claim 9, wherein the human cell is a 293 cell or a PER.C6™ cell.

11. A vaccine composition comprising a replication-defective ChAd vector according to any one of Claims 3-7.

12. An adenoviral E1-expressing human cell comprising the nucleotide sequence set forth in SEQ ID NO:1.

13. An adenoviral E1-expressing human cell comprising the nucleotide sequence set forth in SEQ ID NO: 2.

14. A method of boosting an antigen-specific immune response in a mammal comprising administering to said mammal a sufficient amount of a recombinant ChAd vector comprising a chimpanzee adenovirus genome containing at least a functional deletion of its E1 gene, a nucleotide sequence encoding a target antigen and a promoter sequence capable of directing expression of the nucleotide sequence encoding the target antigen, wherein administration of said chAd vector elicits a boosted response.

15. The method of claim 14, wherein the ChAd vector comprises a complete deletion of its E1 genes and further wherein the vector optionally comprises a deletion of its E3 genes.

16. The method of claim 14, wherein the boosted immune response is specific for an antigen derived from an infectious agent selected from the group consisting of: HIV, HBV, HCV, HPV, HSV1, HSV2, SARS CoV, *Plasmodium malariae*, Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, and *Mycobacterium tuberculosis*.

17. The method of claim 14, wherein the immune response is a boosted immune response that is specific for a TAA.

18. The method of claim 17, wherein the boosted immune response comprises the production of antigen-specific CD8⁺ T cells.

19. The method of claim 14, wherein the boosted immune response comprises the production of antigen-specific CD8⁺ T cells.

20. A method of eliciting an immune response in a naïve mammal comprising administering to said mammal a sufficient amount of a ChAd vector which comprises a chimpanzee adenovirus genome containing at least a functional deletion of its E1 gene, a nucleotide encoding a target antigen and a promoter sequence capable of directing expression of the nucleotide sequence encoding the target antigen, wherein administration of the ChAd vector elicits a primary immune response.

21. The method of claim 20, wherein the primary immune response is specific for an antigen derived from an infectious agent such as, but not limited to HIV, HCV, HPV, HSV1, HSV2, SARS CoV, *Plasmodium malariae*, Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, *Mycobacterium tuberculosis*.

22. The method of claim 14, wherein the immune response is a primary immune response that is specific for a TAA against which the mammal is tolerant.

23. A method of claim according to any one of claims 14 to 22, wherein the recombinant adenovirus comprises a nucleotide sequence encoding a hexon peptide selected from the group consisting of: SEQ ID NOS: 14-21.

5 24. A method of claim according to any one of claims 14 to 22,, wherein the recombinant adenovirus comprises a nucleotide sequence encoding a fiber protein sequence selected from the group consisting of: SEQ ID NOS: 6-15.

10 25. A method of inducing an immune response against an antigen derived from an infectious agent selected from the group consisting of: HIV, HCV, HPV, HSV1, HSV2, SARS, *Plasmodium maleriae*, Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, and *Mycobacterium tuberculosis* comprising: (a) priming a host to respond to a infectious agent-antigen by administering a first vaccine composition comprising a nucleotide sequence encoding a infectious agent-antigen against which an antigen-specific immune response is
15 desired; and (b) boosting the immune response of step (a) by administering a second vaccine composition comprising a recombinant ChAd vector containing at least a functional deletion of its E1 gene, and in the site of the E1 gene deletion, a sequence comprising a promoter capable of directing expression of DNA encoding the same infectious agent-antigen delivered in the priming step,
20 wherein administration of the boosting composition elicits an immune response which has the effect of conferring protective immunity.

25 26. The method according to claim 25, wherein the first vaccine composition comprises plasmid DNA which is administered intramuscularly in combination with electrical stimulation.

30 27. The method of claim 25, wherein the second vaccine composition comprises a ChAd vector comprising DNA encoding an antigen derived from an infectious agent selected from the group consisting of: HIV, HCV, HPV, HSV1, HSV2, SARS, Malaria, Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, and *Mycobacterium tuberculosis*.

28. The method of Claim 27, wherein the immune response comprises the production of antigen-specific CD8+ T cells.

29. The method of claim 27, wherein the ChAd vector is selected from the group consisting of: ChAd3, ChAd6, ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd 16, ChAd17, and ChAd19.
- 5 30. The method of claim 25, wherein the ChAd vector comprises a nucleotide sequence encoding a hexon peptide selected from the group consisting of: SEQ ID NO: 16-25.
31. The method of claim 25, wherein the ChAd vector comprises a nucleotide sequence encoding a fiber peptide selected from the group consisting of: SEQ ID NOS: 6-15.
- 10 32. The method of claim 25, wherein the first and second vaccine compositions are both ChAd vectors characterized by different serotypes.
33. A method of breaking host tolerance to a self-antigen comprising: (a) priming a host to respond to a self-antigen by administering a first vaccine composition comprising a nucleotide sequence encoding a self-antigen against which an antigen-specific immune response is desired, thereby eliciting a primed response; and (b) boosting the primed immune response of step (a) by administering a second vaccine composition comprising a recombinant ChAd vector containing at least a functional deletion of its E1 gene, and in the site of the E1 gene deletion, a sequence comprising a promoter capable of directing expression of DNA encoding the same self-antigen delivered in the priming step, wherein administration of the boosting composition elicits an immune response which has the effect of breaking host tolerance to the self-antigen.
- 15 34. The method according to claim 33, wherein the first vaccine composition comprises plasmid DNA which is administered intramuscularly in combination with electrical stimulation.
- 20 35. The method of claim 33, wherein the second vaccine composition comprises a ChAd vector comprising DNA encoding a self antigen selected from the group consisting of: HER2 NEU, CEA, HEPCAM, PSA, PSMA, TELOMERASE, GP100, MELAN-A/MART-1, MUC-1, NY-ESO-1, SURVIVIN, STROMELYSIN 3, TYROSINASE, MAGE3, CML68, CML66, OY-25 TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, HKLP2, 5T4 and VEGFR2.
- 30 36. The method of Claim 35, wherein the immune response comprises the production of antigen-specific CD8+ T cells.
- 35

37. The method of claim 35, wherein the ChAd vector is selected from the group consisting of: ChAd3, ChAd6, ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16,
5 ChAd17, and ChAd19.

38. The method of claim 35, wherein the ChAd vector comprises a nucleotide sequence encoding a hexon peptide selected from the group consisting of: SEQ ID NO: 16-25.

10 39. The method of claim 35, wherein the ChAd vector comprises a nucleotide sequence encoding a fiber peptide selected from the group consisting of: SEQ ID NOS: 6-15.

40. The method of claim 35, wherein the first and second vaccine compositions are both ChAd vectors characterized by different serotypes.
15

41. The method of claim 35, wherein the host is a tumor-bearing mammal who has developed resistance to cancer chemotherapy.
20

20

25

30

35

5

10

ABSTRACT OF THE DISCLOSURE

The present invention provides recombinant replication-defective adenoviral vectors derived from chimpanzee adenoviruses and methods for generating recombinant
15 adenoviruses in human E1-expressing cell lines. The invention also provides compositions and methods suitable for use for the delivery and expression of transgenes encoding immunogens against which a boosted immune response is desired. The invention further provides methods of generating clinical grade vector stocks suitable for use in humans. In a particular embodiment
the invention contemplates the use of vectors comprising transgenes which encode tumor
20 associated antigens in vaccines and pharmaceutical compositions for the prevention and treatment of cancer.

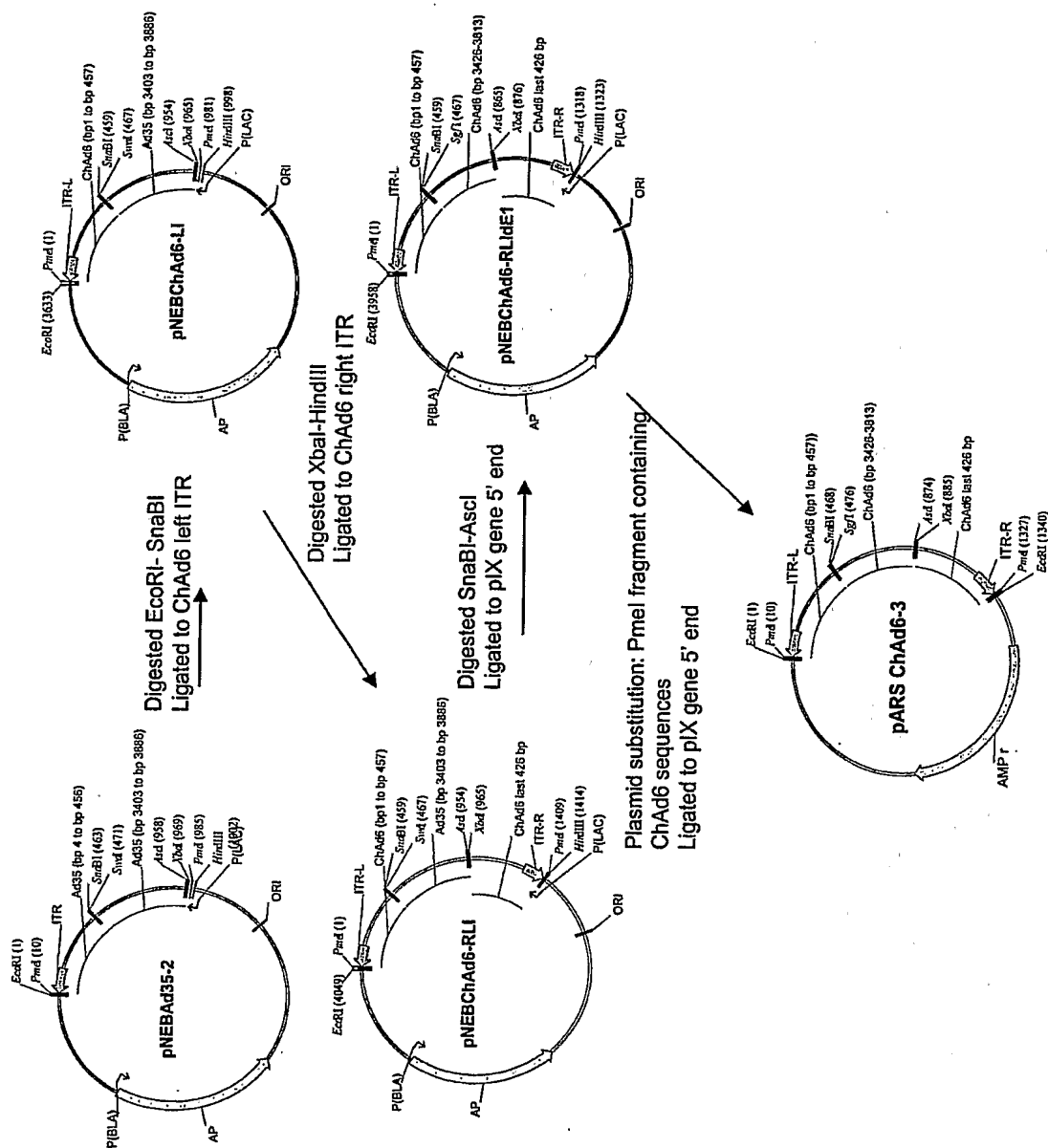


Fig. 1

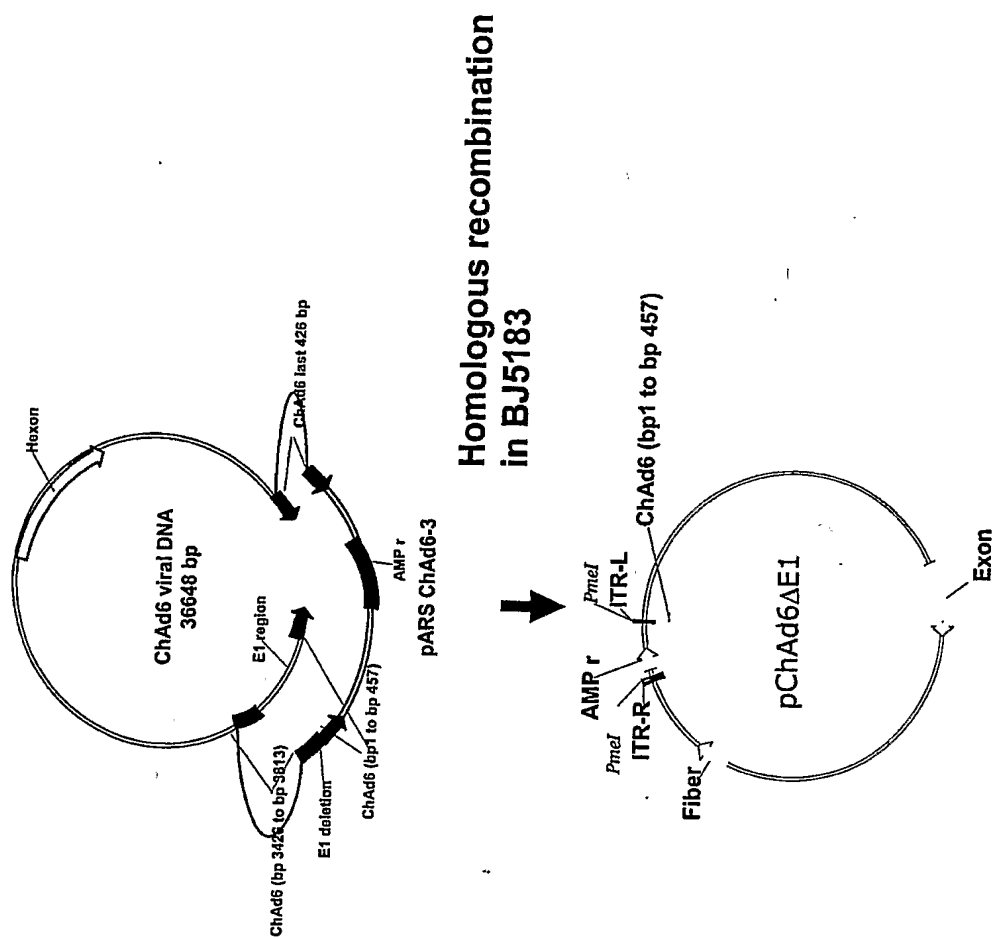


Fig. 2

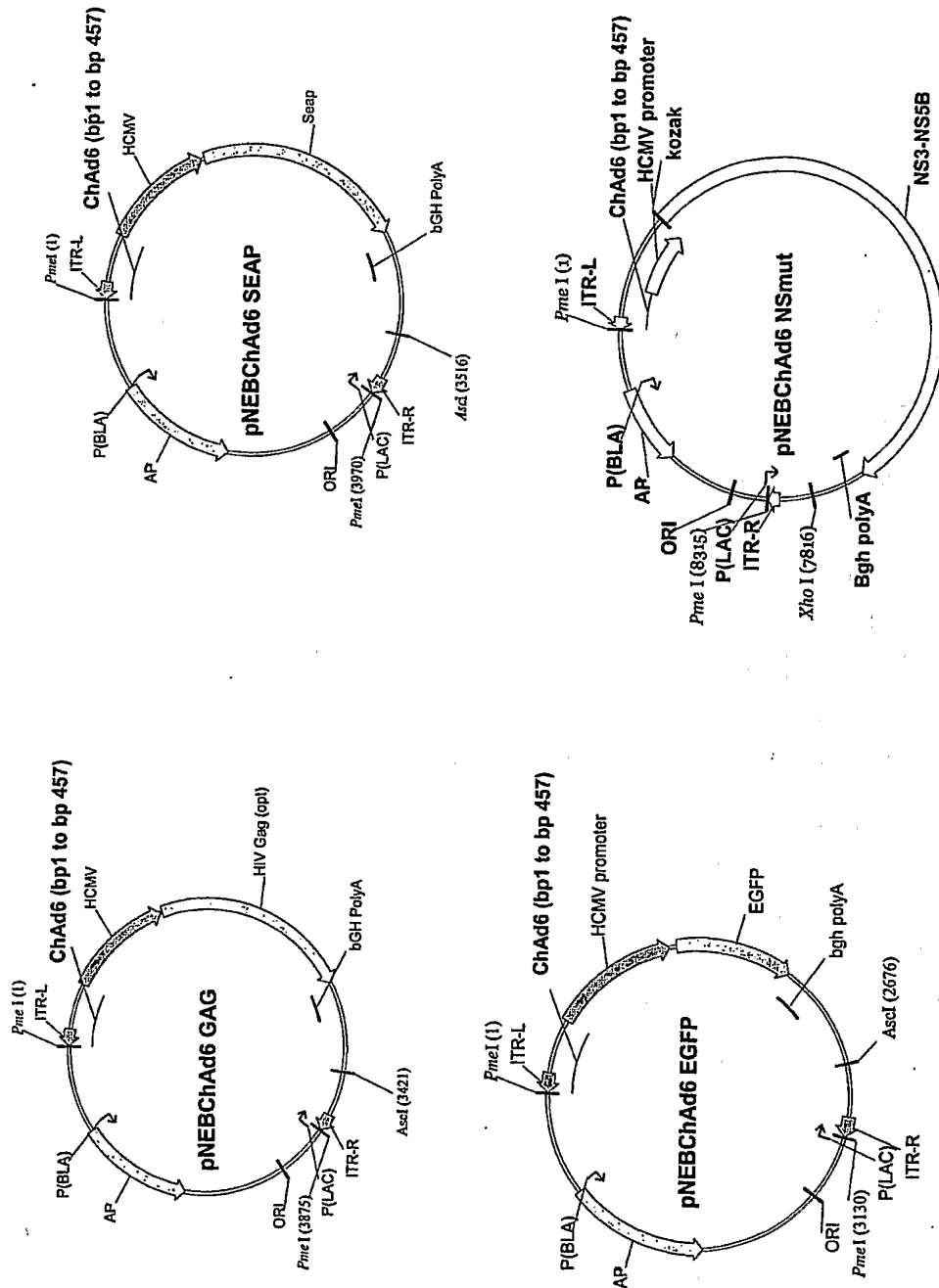


Fig. 3

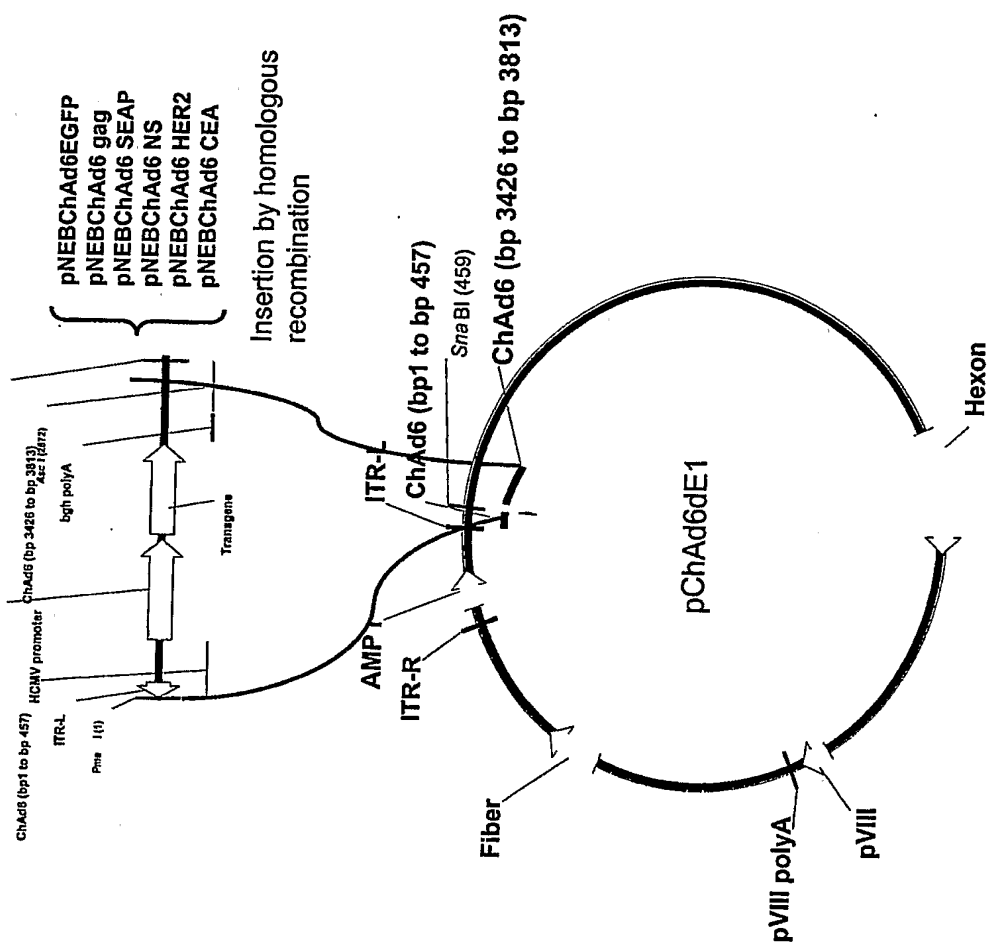


Fig. 4

SEQ ID NO:1

5/153

1 CATCATCAAT AATATACCTT ATTTTGGATT GAAGCCAATA TGATAATGAG ATGGGCGGCG
61 CGAGGCGGGG CGCGGGGCGG GAGGCGGGTT TGGGGGCGGG CCGGCGGGCG GGGCGGTGTG
121 GCGGAAGTGG ACTTTGTAAG TGTGGCGGAT GTGACTTGCT AGTGCCGGGC GCGGTAAAAG
181 TGACGTTTTTC CGTGCGCGAC AACGCCCCCG GGAAGTGACA TTTTCCCCGC GGTTTTTTACC
241 GGATGTTGTA GTGAATTTGG GCGTAACCAA GTAAGATTTG GCCATTTTCG CGGGAAAACT
301 GAAACGGGGA AGTGAAATCT GATTAATTTT GCGTTAGTCA TACCGCGTAA TATTGTCTA
361 GGGCCGAGGG ACTTTGGCCG ATTACGTGGA GGAATCGCCC AGGTGTTTTT TGAGGTGAAT
421 TTCCGCGTTC CGGGTCAAAG TCTCCGTTTT ATTATTATAG TCAGCTGACG CGGAGTGTAT
481 TTATACCCTC TGATCTCGTC AAGAGGCCAC TCTTGAGTGC CAGCGAGTAG AGTTTTCTCC
541 TCTGCCGCTC TCCGCTCCGC TCCGCTCGGC TCTGACACCG GGGAAAAAAT GAGACATTTT
601 ACCTACGATG GCGGTGTGCT CACCGGCCAG CTGGCTGCTG AGGTCCTGGA CACCCTGATC
661 GAGGAGGTAT TGGCCGATAA TTATCCTCCC TCGACTCCTT TTGAGCCACC TACACTTCAC
721 GAACTATACG ATCTGGATGT GGTGGGGCCC AGCGATCCGA ACGAGCAGGC GGTTCCTAGT
781 TTTTTCCTCAG AGTCCATGTT GTTGGCCAGC CAGGAGGGGG TCGAACTTGA GACCCCTCCT
841 CCGATCGTGG ATTCCCCCGA TCCGCCGCGAG CTGACTAGGC AGCCCGAGCG CTGTGCGGGA
901 CCTGAGACTA TGCCCCAGCT GCTACCTGAG GTGATCGATC TCACCTGTAA TGAGTCTGGT
961 TTTCCACCCA GCGAGGATGA GGACGAAGAG GGTGAGCAGT TTGTGTTAGA TTCTGTGGAA
1021 CAACCCGGGC GAGGATGCAG GTCTTGTCAG TATCACCAGA AAAACACAGG AGACTCCAG
1081 ATTATGTGTT CTCTGTGTTA TATGAAGATG ACCTGTATGT TTATTTACAG TAAGTTTATC
1141 ATCGGTGGGC AGGTGGGCTA TAGTGTGGGT GGTGGTCTTT GGGGGGTTTT TTAATATATG
1201 TCAGGGGTTA TGCTGAAGAC TTTTATTATTG TGATTTTAA AGGTCCAGTG TCTGAGCCCC
1261 AGCAAGAACC TGAACCGGAG CCTGAGCCTT CTCGCCCCAG GAGAAAGCCT GTAATCTTAA
1321 CTAGACCCAG CGCACCAGTA GCGAGAGGCC TCAGCAGCGC GGAGACCACC GACTCCGGTG
1381 CTTCCTCATC ACCCCCGGAG ATTCACCCCC TGGTGCCCCCT ATGTCCCGTT AAGCCCGTTG
1441 CCGTGAGAGT CAGTGGGCGG CGGTCTGCTG TGGAGTGCAT TGAGGACTTG CTTTTTGATT
1501 CACAGGAACC TTTGGACTTG AGCTTGAAAC GCCCCAGGCA TTAAACCTGG TCACCTGGAC
1561 TGAATGAGTT GACGCCTATG TTTGCTTTTG AATGACTTAA TGTGTATAGA TAATAAGAG
1621 TGAGATAATG TTTTAATTGC ATGGTGTGTT TAACTTGGGC GGAGTCTGCT GGGTATATAA
1681 GCTTCCCTGG GCTAAACTTG GTTACACTTG ACCTCATGGA GGCCTGGGAG TGTTTGAGAA

Fig. 5A

SEQ ID NO:1

6/153

1741 ACTTTGCCGG AGTTCGTGCC TTGCTGGACG AGAGCTCTAA CAATACCTCT TGGTGGTGGG
1801 GGTATTTGTG GGGCTCTCCC CAGGGCAAGT TAGTTTGTAG AATCAAGGAG GATTACAAGT
1861 GGGAATTTGA AGAGCTTTTG AAATCCTGTG GTGAGCTATT GGATTCTTTG AATCTAGGCC
1921 ACCAGGCTCT CTTCCAGGAG AAGGTCATCA GGACTTTGGA TTTTCCACA CCGGGGCGCA
1981 TTGCAGCCGC GGTGCTTTT CTAGCTTTTT TGAAGGATAG ATGGAGCGAA GAGACCCACT
2041 TGAGTTCGGG CTACGTCTTG GATTTTCTGG CCATGCAACT GTGGAGAGCA TGGATCAGAC
2101 ACAAGAACAG GCTGCAACTG TTGTCTTCCG TCCGCCCCGT GCTGATTCCG GCGGAGGAGC
2161 AACAGGCCGG GTCAGAGGAC CGGGCCCCGT GGGATCCGGA GGAGAGGGCA CCGAGGCCGG
2221 GCGAGAGGAG CGCGCTGAAC CTGGGAACCG GGCTGAGCGG CCATCCACAT CGGGAGTGAA
2281 TGTCGGGCAG GTGGTGGATC TTTTCCAGA ACTGCGGCGG ATTTTACTA TTAGGGAGGA
2341 TGGGCAATTT GTTAAGGGTC TTAAGAGGGA GAGGGGGGCT TCTGAGCATA ACGAGGAGGC
2401 CAGTAATTTA GCTTTTAGCT TGATGACCAG ACACCGTCCA GAGTGCATCA CTTTTCAGCA
2461 GATTAAGGAC AATTGTGCCA ATGAGTTGGA TCTGTTGGGT CAGAAGTATA GCATAGAGCA
2521 GCTGACCACT TACTGGCTGC AGCCGGGTGA TGATCTGGAG GAAGCTATTA GGGTGTATGC
2581 TAAGGTGGCC CTGCGGCCCC ATTGCAAGTA CAAGCTCAAG GGGCTGGTGA ATATCAGGAA
2641 TTGTTGCTAC ATTTCTGGCA ACGGGGCGGA GGTGGAGATA GAGACCGAAG ACAGGGTGGC
2701 TTTGAGATGC AGCATGATGA ATATGTGGCC GGGGGTGCTG GGCATGGACG GGGTGGTGAT
2761 TATGAATGTG AGGTTACGG GGGCCAAC TT TAACGGCACG GTGTTTTTGG GGAACACCAA
2821 CCTGGTCTTG CACGGGGTGA GCTTCTATGG GTTTAACAAC ACCTGTGTGG AGGCCTGGAC
2881 CGATGTGAAG GTCCGCGGTT GCGCCTTTTA TGGATGTTGG AAGGCCATAG TGAGCCGCCC
2941 TAAGAGCAGG AGTTCCATTA AGAAATGCTT GTTTGAGAGG TGCACCTTGG GGATCCTGGC
3001 CGAGGGCAAC TGCAGGGTGC GCCACAATGT GGCCTCCGAG TCGGTTGCT TCATGCTAGT
3061 CAAGAGCGTG GCGGTAATCA AGCATAATAT GGTGTGCGGC AACAGCGAGG ACAAGGCCTC
3121 ACAGATGCTG ACCTGCACGG ATGGCAACTG CCACTTGCTG AAGACCATCC ATGTAACCAG
3181 CCACAGCCGG AAGGCCTGGC CCGTGTTGCA GCACAACTTG CTGACCCGCT GCTCCTTGCA
3241 TCTGGGCAAC AGGCGGGGGG TGTTCTGCGC CTATCAATGC AACTTTAGTC ACACCAAGAT
3301 CTTGCTAGAG CCCGAGAGCA TGTCCAAGGT GAACTTGAAC GGGGTGTTT ACATGACCAT
3361 GAAGATCTGG AAGGTGCTGA GGTACGACGA GACCAGGTCC CGGTGCAGAC CCTGCGAGTG
3421 CGGGGGCAAG CATATGAGGA ACCAGCCCGT GATGCTGGAT GTGACCGAGG AGCTGAGGAC

Fig. 5B

SEQ ID NO:1

7/153

3481 AGACCACTTG GTTCTGGCCT GCACCAGGGC CGAGTTTGGT TCTAGCGATG AAGACACAGA
3541 TTGAGGTGGG TGAGTGGGCG TGGCCTGGGG TGGTCATGAA AATATATAAG TTGGGGGTCT
3601 TAGGGTCTCT TTATTTGTGT TGCAGAGACC GCCGGAGCCA TGAGCGGGAG CAGCAGCAGC
3661 AGCAGTAGCA GCAGCGCCTT GGATGGCAGC ATCGTGAGCC CTTATTTGAC GACGCGGATG
3721 CCCCCTGGG CCGGGGTGCG TCAGAATGTG ATGGGCTCCA GCATCGACGG CCGACCCGTC
3781 CTGCCCCGAA ATTCCGCCAC GCTGACCTAT GCGACCGTCG CGGGGACGCC GTTGACGCC
3841 ACCGCGCCG CCGCCGCCAC CGCAGCCGCC TCGGCCGTGC GCAGCCTGGC CACGGACTTT
3901 GCATTCTTG GACCACTGGC GACAGGGGCT ACTTCTCGGG CCGCTGCTGC CGCCGTTGCG
3961 GATGACAAGC TGACCGCCCT GCTGGCGCAG TTGGATGCGC TTACTCGGGA ACTGGGTGAC
4021 CTTTCTCAGC AGGTCATGGC CCTGCGCCAG CAGGTCTCCT CCCTGCAAGC TGGCGGGAAT
4081 GCTTCTCCA CAAATGCCGT TTAAGATAAA TAAAACCAGA CTCTGTTTGG ATTAAAGAAA
4141 AGTAGCAAGT GCATTGCTCT CTTTATTTCA TAATTTTCCG CGCGCGATAG GCCCTAGACC
4201 AGCGTTCTCG GTCGTTGAGG GTGCGGTGTA TCTTCTCCAG GACGTGGTAG AGGTGGCTCT
4261 GGACGTTGAG ATACATGGGC ATGAGCCCGT CCCGGGGGTG GAGGTAGCAC CACTGCAGAG
4321 CTTTCATGCTC CGGGGTGGTG TTGTAGATGA TCCAGTCGTA GCAGGAGCGC TGGGCATGGT
4381 GCCTAAAAAT GTCCTTCAGC AGCAGGCCGA TGGCCAGGGG GAGGCCCTTG GTGTAAGTGT
4441 TTACAAAACG GTTAAGTTGG GAAGGGTGCA TTCGGGGAGA GATGATGTGC ATCTTGGACT
4501 GTATTTTATG ATTGGCGATG TTTCCGCCCA GATCCCTTCT GGGATTTCATG TTGTGCAGGA
4561 CCACCAGTAC AGTGTATCCG GTGCACTTGG GGAATTTGTC ATGCAGCTTA GAGGGAAAAG
4621 CGTGGAAGAA CTTGGAGACG CCCTTGTTGGC CTCCCAGATT TTCCATGCAT TCGTCCATGA
4681 TGATGGCAAT GGGCCCGCGG GAGGCAGCTT GGGCAAAGAT ATTTCTGGGG TCGCTGACGT
4741 CGTAGTTGTG TTCCAGGGTG AGGTGCTCAT AGGCCATTTT TACAAAGCGC GGGCGGAGGG
4801 TGCCCGACTG GGGGATGATG GTCCCCTCTG GCCCTGGGGC GTAGTTGCCC TCGCAGATCT
4861 GCATTTCCA GGCCTTAATC TCGGAGGGGG GAATCATATC CACCTGCGGG GCGATGAAGA
4921 AAACGGTTTC CGGAGCCGGG GAGATTAAC TGGATGAGAG CAGGTTTCTA AGCAGCTGTG
4981 ATTTTCCACA ACCGGTGGGC CCATAAATAA CACCTATAAC CGGTTGCAGC TGGTAGTTTA
5041 GAGAGCTGCA GCTGCCGTCG TCCCGGAGGA GGGGGGCCAC CTCGTTGAGC ATGTCCCTGA
5101 CGCGCATGTT CTCCCCGACC AGATCCGCCA GAAGGCGCTC GCCGCCAGG GACAGCAGCT
5161 CTTGCAAGGA AGCAAAGTTT TTCAGCGGCT TGAGGCCGTC CGCCGTGGGC ATGTTTTTCA

Fig. 5C

SEQ ID NO:1

8/153

5221 GGGTCTGGCT CAGCAGCTCC AGGCGGTCCC AGAGCTCGGT GACGTGCTCT ACGGCATCTC
5281 TATCCAGCAT ATCTCCTCGT TTCGCGGGTT GGGGCGACTT TCGCTGTAGG GCACCAAGCG
5341 GTGGTCGTCC AGCGGGGCCA AAGTCATGTC CTTCCATGGG CGCAGGGTCC TCGTCAGGGT
5401 GGTCTGGGTC ACGGTGAAGG GGTGCGCTCC GGGCTGAGCG CTTGCCAAGG TGCGCTTGAG
5461 GCTGGTTCTG CTGGTGCTGA AGCGCTGCCG GTCTTCGCCC TGCGCGTCGG CCAGGTAGCA
5521 TTTGACCATG GTGTCATAGT CCAGCCCCTC CGCGGCGTGT CCCTTGCGCG GCAGCTTGCC
5581 CTTGGAGGTG GCGCCGCACG AGGGGCAGAG CAGGCTCTTG AGCGCGTAGA GCTTGGGGGC
5641 GAGGAAGACC GATTGCGGGG AGTAGGCGTC CGCGCCGCAG ACCCCGCACA CCGTCTCGCA
5701 CTCCACCAGC CAGGTGAGCT CGGGGCGCGC CGGGTCAAAA ACCAGGTTTC CCCCATGCTT
5761 TTTGATGCGT TTCTTACCTC GGGTCTCCAT GAGGTGGTGT CCCCCTCGG TGACGAAGAG
5821 GCTGTCCGTG TCTCCGTAGA CCGACTTGAG GGGTCTTTTC TCCAGGGGGG TCCCTCGGTC
5881 TTCTCTGTAG AGGAATCGG ACCACTCTGA GACGAAGGCC CGCGTCCAGG CCAGGACGAA
5941 GGAGGCTATG TGGGAGGGGT AGCGGTCGTT GTCCACTAGG GGGTCCACCT TCTCCAAGGT
6001 GTGAAGACAC ATGTCGCCTT CCTCGGCGTC CAGGAAGGTG ATTGGCTTGT AGGTGTAGGC
6061 CACGTGACCG GGGGTTCCTG ACGGGGGGGT ATAAAAGGGG GTGGGGGCGC GCTCGTCGTC
6121 ACTCTCTTCC GCATCGCTGT CTGCGAGGGC CAGCTGCTGG GGTGAGTATT CCTCTCGAA
6181 GGCGGGCATG ACCTCCGCGC TGAGGTTGTC AGTTTCCTAAA AACGAGGAGG ATTTGATGTT
6241 CACCTGTCCC GAGGTGATAC CTTTGAGGGT ACCCGCGTCC ATCTGGTCAG AAAACACGAT
6301 CTTTTTATTG TCCAGCTTGG TGGCGAACGA CCCGTAGAGG GCGTTGGAGA GCAGCTTGCC
6361 GATGGAGCGC AGGGTCTGGT TCTTGTCCTT GTCGGCGCGC TCCTTGGCCG CGATGTTGAG
6421 CTGCACGTAC TCGCGCGCGA CGCAGCGCCA CTCGGGGAAG ACGGTGGTGC GCTCGTCGGG
6481 CACCAGGCGC ACGCGCCAGC CGCGGTTGTG CAGGGTGACC AGGTCCACGC TGGTGGCGAC
6541 CTCGCCGCGC AGGCGCTCGT TGGTCCAGCA GAGACGGCCG CCCTTGCGCG AGCAGAAGGG
6601 GGGCAGGGGG TCGAGCTGGG TCTCGTCCGG GGGGTCCGCG TCCACGGTGA AAACCCCGG
6661 GCGCAGGCGC GCGTCGAAGT AGTCTATCTT GCAACCTTGC ATGTCCAGCG CCTGCTGCCA
6721 GTCGCGGGCG GCGAGCGCGC GCTCGTAGGG GTTGAGCGGC GGGCCCCAGG GCATGGGGTG
6781 GGTGAGTGCG GAGGCGTACA TGCCGCAGAT GTCATAGACG TAGAGGGGCT CCCGCAGGAC
6841 CCCGATGTAG GTGGGGTAGC AGCGGCCGCC GCGGATGCTG GCGCGCACGT AGTCATACAG
6901 CTCGTGCGAG GGGGCGAGGA GGTGCGGGCC CAGGTTGGTG CGGGCGGGGC GCTCCGCGCG

Fig. 5D

SEQ ID NO:1

9/153

6961 GAAGACGATC TGCCTGAAGA TGGCATGCGA GTTGAAGAG ATGGTGGGGC GCTGGAAGAC
7021 GTTGAAGCTG GCGTCCTGCA GGCCGACGGC GTCGCGCACG AAGGAGGCGT AGGAGTCGCG
7081 CAGCTTGTGT ACCAGCTCGG CCGTGACCTG CACGTCGAGC GCGCAGTAGT CGAGGGTCTC
7141 GCGGATGATG TCATATTTAG CCTGCCCTT CTTTTTCCAC AGCTCGCGGT TGAGGACAAA
7201 CTCTTCGCGG TCTTTCCAGT ACTCTTGAT CGGGAAACCG TCCGGTTCCG AACGTAAGA
7261 GCCTAGCATG TAGAACTGGT TGACGGCCTG GTAGGCGCAG CAGCCCTTCT CCACGGGGAG
7321 GGCGTAGGCC TGC GCGGCCT TGC GGAGCGA GGTGTGGGTC AGGGCGAAGG TGTCCCTGAC
7381 CATGACTTTG AGGTACTGGT GCTTGAAGTC GGAGTCGTCG CAGCCGCCCC GCTCCCAGAG
7441 CGAGAAGTCG GTGCGCTTCT TGAGCGGGG GTTGGGCAGA GCGAAGGTGA CATCGTTGAA
7501 GAGGATTTTG CCCGCGCGGG GCATGAAGTT GCGGGTGATG CGGAAGGGCC CCGGCACTTC
7561 AGAGCGGTTG TTGATGACCT GGGCGGCGAG CACGATCTCG TCGAAGCCGT TGATGTTGTG
7621 GCCCACGATG TAGAGTTCCA GGAAGCGGGG CCGGCCCTTT ACGGTGGGCA GCTTCTTTAG
7681 CTCTTCGTAG GTGAGCTCCT CGGGCGAGGC GAGGCCGTGC TCGGCCAGGG CCCAGTCCGC
7741 GAGGTGCGGG TTGTCTCTGA GGAAGGACTC CCAGAGGTCG CGGGCCAGGA GGGTCTGCAG
7801 GCGGTCCCTG AAGGTCTGA ACTGGCGGCC CACGGCCATT TTTTCGGGGG TGATGCAGTA
7861 GAAGGTGAGG GGGTCTTGCT GCCAGCGGTC CCAGTCGAGC TGCAGGGCGA GGTGCGCGCG
7921 GGCGGTGACC AGGCGCTCGT CGCCCCGAA TTTCATGACC AGCATGAAGG GCACGAGCTG
7981 CTTTCCGAAG GCCCCATCC AAGTGTAGGT CTCTACATCG TAGGTGACAA AGAGGCGCTC
8041 CGTGCGAGGA TGCGAGCCGA TCGGGAAGAA CTGGATCTCC CGCCACCACT TGGAGGAGTG
8101 GCTGTTGATG TGGTGGAAGT AGAAGTCCCG TCGCCGGGCC GAACACTCGT GCTGGCTTTT
8161 GTAAAAGCGA GCGCAGTACT GGCAGCGCTG CACGGGCTGT ACCTCCTGCA CGAGATGCAC
8221 CTTTCGCCCC CGCACGAGGA AGCCGAGGGG AAATCTGAGC CCCCCGCTG GCTCGCGGCA
8281 TGGCTGGTGC TCTTCTACTT TGGATGCGTG TCCGTCTCCG TCTGGCTCCT CGAGGGGTGT
8341 TACGGTGGAG CGGACCACCA CGCCGCGCGA GCCGAGGTC CAGATATCGG CGCGCGGCGG
8401 TCGGAGTTTG ATGACGACAT CGCGCAGCTG GGAGCTGTCC ATGGTCTGGA GCTCCCGCGG
8461 CGGCGGCAGG TCAGCCGGGA GTTCTTGAG GTTCACCTCG CAGAGTCGGG CCAGGGCGCG
8521 GGGCAGGTCT AGGTGGTACC TGATCTCTAG GGGCGTGTG GTGGCGGCGT CGATGGCTTG
8581 CAGGAGCCCG CATCCCCGGG GGGCGACGAC GGTGCCCCGC GGGGTGGTGG TGGTGGTGGT
8641 GGTGGTGGTG GTGGCGGTGC AGCTCAGAAG CGGTGCCGCG GCGGGGCCCC CGGAGGTAGG

Fig. 5E

SEQ ID NO:1

10/153

8701 GGGGGCTCCG GTCCCCCGCG CAGGGGCGGC AGCGGCACGT CGCGGTGGAG CGCGGGCAGG
8761 AGTTGGTGCT GTCCCCGAG GTTGCTGGCG AAGGCGACGA CGCGCGGTT GATCTCCTGG
8821 ATCTGGCGCC TCTGCGTGAA GACGACGGGC CCGGTGAGCT TGAACCTGAA AGAGAGTTCC
8881 ACAGAATCAA TCTCGGTGTC ATTGACCGCG GCCTGGCGCA GGATCTCCTG CACGTCTCCC
8941 GAGTTGTCTT GGTAGGCGAT CTCGGCCATG AACTGCTCGA TCTCTTCTC CTGGAGGTCT
9001 CCGCGTCCGG CGCGTTCCAC GGTGGCCGCC AGGTCGTTGG AGATGCGCCC CATGAGCTGC
9061 GAGAAGGCGT TGAGTCCGCC CTCGTCCAG ACTCGGCTGT AGACCACGCC CCCCTGGTCA
9121 TCGCGGGCGC GCATGACCAC CTGCGCGAGG TTGAGCTCCA CGTGCCGCGC GAAGACGGCG
9181 TAGTTGCGCA GACGCTGGAA GAGGTAGTTG AGGGTGGTGG CGGTGTGCTC GGCCACGAAG
9241 AAGTTCATGA CCCAGCGGCG CAACGTGGAT TCGTTGATGT CCCCCAAGGC CTCCAGCCGT
9301 TCCATGGCCT CGTAGAAGTC CACGGCGAAG TTGAAAACT GGGAGTTGCG CGCCGACACG
9361 GTCAACTCCT CCTCCAGAAG ACGGATGAGC TCGGCGACGG TGTGCGGCAC CTCGCGCTCG
9421 AAGGCTATGG GGATCTCTTC CTCCGCTAGC ATCACCACCT CCTCCTCTTC CTCTCTTCT
9481 GGCAC TTCCA TGATGGCTTC CTCCTCTTCG GGGGGCGGCG GCGGCGGCGG TGGGGGAGGG
9541 GGCCTCTGCG GCCGGCGGCG GCGCACCGGG AGGCGGTCCA CGAAGCGCGC GATCATCTCC
9601 CCGCGGCGGC GCGCATGGT CTCGGTGACG GCGGCGCCGT TCTCCCGGGG GCGCAGTTGG
9661 AAGACGCCGC CGGACATCTG GTGCTGGGGC GGGTGGCCGT GAGGCAGCGA AACGGCGCTG
9721 ACGATGCATC TCAACAATTG CTGCGTAGGT ACGCCGCCGA GGGACCTGAG GGAGTCCATA
9781 TCCACCGGAT CCGAAAACCT TTCGAGGAAG GCGTCTAACC AGTCGCAGTC GCAAGGTAGG
9841 CTGAGCACCG TGGCGGGCGG CGGGGGGTGG GGGGAGTGTC TGGCGGAGGT GCTGCTGATG
9901 ATGTAATTGA AGTAGGCGGA CTGACACGG CGGATGGTCG ACAGGAGCAC CATGTCCTTG
9961 GGTCCGGCCT GCTGGATGCG GAGGCGGTCG GCTATGCCCC AGGCTTCGTT CTGGCATCGG
10021 CGCAGGTCCT TG TAGTAGTC TTGCATGAGC CTTTCCACCG GCACCTCTTC TCCTTCTCT
10081 TCTGCTTCTT CCATGTCTGC TTCGGCCCTG GGGCGGCGCC GCGCCCCCT GCCCCCCATG
10141 CGCGTGACCC CGAACCCCT GAGCGGTTGG AGCAGGGCCA GGTCGGCGAC GACGCGCTCG
10201 GCCAGGATGG CCTGCTGCAC CTGCGTGAGG GTGGTTTGA AGTCATCAA GTCCACGAAG
10261 CGGTGGTAGG CGCCCGTGTT GATGGTGTAG GTGCAGTTGG CCATGACGGA CCAGTTGACG
10321 GTCTGGTGGC CCGGTTGCGA CATCTCGGTG TACCTGAGTC GCGAGTAGGC GCGGGAGTCG
10381 AAGACGTAGT CGTTGCAAGT CCGCACCAGG TACTGGTAGC CCACCAGGAA GTGCGGCGGC

Fig. 5F

SEQ ID NO:1

11/153

10441 GGCTGGCGGT AGAGGGGCCA GCGCAGGGTG GCGGGGGCTC CGGGGGCCAG GTCTTCCAGC
10501 ATGAGGCGGT GGTAGGCGTA GATGTACCTG GACATCCAGG TGATACCCGC GGCGGTGGTG
10561 GAGGCGCGCG GGAAGTCGCG CACCCGGTTC CAGATGTTGC GCAGGGGCAG AAAGTGCTCC
10621 ATGGTAGGCG TGCTCTGTCC AGTCAGACGC GCGCAGTCGT TGATACTCTA GACCAGGGAA
10681 AACGAAAGCC GGTACAGCGG CACTCTTCCG TGGTCTGGTG AATAGATCGC AAGGGTATCA
10741 TGGCGGAGGG CCTCGGTTCC AGCCCCGGGT CCGGGCCGGA CGGTCCGCCA TGATCCACGC
10801 GGTACCGCC CGCGTGTCGA ACCCAGGTGT GCGACGTCAG ACAACGGTGG AGTGTTCTTT
10861 TTGGCGTTTT TCTGGCCGGG CGCCGGCGTC GCGTAAGAGA CTAAGCCGCG AAAGCGAAAG
10921 CAGTAAGTGG CTCGCTCCCC GTAGCCGGAG GGATCCTTGC TAAGGGTTGC GTTGCGGCGA
10981 ACCCCGGTTC GAATCCCGTA CTCGGGCCGG CCGGACCCGC GGCTAAGGTG TTGGATTGGC
11041 CTCCCCCTCG TATAAAGACC CCGCTTGCGG ATTGACTCCG GACACGGGGA CGAGCCCCCTT
11101 TTATTTTTGC TTTCCCCAGA TGATCCGGT GCTGCGGCAG ATGCGCCCCC CGCCCCAGCA
11161 GCAGCAACAA CACCAGCAAG AGCGGCAGCA ACAGCAGCGG GAGTCATGCA GGGCCCCCTC
11221 ACCCACCTC GCGGGGCCGG CCACCTCGGC GTCCGCGGCC GTGTCTGGCG CCTGCGGCGG
11281 CGGCGGGGGG CCGGCTGACG ACCCCGAGGA GCGCCGCGG CGCAGGGCCA GACACTACCT
11341 GGACCTGGAG GAGGGCGAGG GCCTGGCGCG GCTGGGGGCG CCGTCTCCCG AGCGCCACCC
11401 GCGGGTGACG CTGAAGCGCG ACTCGCGCGA GCGGTACGTG CCTCGGCAGA ACCTGTTACG
11461 GGACCGCGCG GCGGAGGAGC CCGAGGAGAT GCGGGACAGG AGGTTACGCG CAGGGCGGGA
11521 GCTGCGGCAG GGGCTGAACC GCGAGCGGCT GCTGCGCGAG GAGGACTTTG AGCCCGACGC
11581 GCGGACGGGG ATCAGCCCCG CGCGCGCGCA CGTGGCGGCC GCCGACCTGG TGACGGCGTA
11641 CGAGCAGACG GTGAACCAGG AGATCAACTT CAAAAGAGT TTCAACAACC ACGTGCGCAC
11701 GCTGGTGGCG CGCGAGGAGG TGACCATCGG GCTGATGCAC CTGTGGGACT TTGTAAGCGC
11761 GCTGGTGACG AACCCCAACA GCAAGCCTCT GACGGCGCAG CTGTTCTGA TAGTGCAGCA
11821 CAGCAGGGAC AACGAGGCGT TTAGGGACGC GCTGCTGAAC ATCACCAGC CCGAGGGTGC
11881 GTGGCTGCTG GACCTGATTA ACATCCTGCA GAGCATAGTG GTGCAGGAGC GCAGCCTGAG
11941 CCTGGCCGAC AAGGTGGCGG CCATCAACTA CTCGATGCTG AGCCTGGGCA AGTTTTACGC
12001 GCGCAAGATC TACCAGACGC CGTACGTGCC CATAGACAAG GAGGTGAAGA TCGACGGTTT
12061 TTACATGCGC ATGGCGCTGA AGGTGCTCAC CCTGAGCGAC GACCTGGGCG TGTACCGCAA
12121 CGAGCGCATC CACAAGGCCG TGAGCGTGAG CCGGCGGCGC GAGCTGAGCG ACCGCGAGCT

Fig. 5G

SEQ ID NO:1

12/153

12181 GATGCACAGC CTGCAGCGGG CGCTGGCGGG CGCCGGCAGC GCGACAGGG AGGCGGAGTC
12241 CTACTTCGAT GCGGGGGCGG ACCTGCGCTG GCGCCCCAGC CGGCGGGCCC TGGAGGCCGC
12301 GGGGGTCCGC GAGGACTATG ACGAGGACGG CGAGGAGGAT GAGGAGTACG AGCTAGAGGA
12361 GGGCGAGTAC CTGGACTAAA CCGCGGGTGG TGTTCCTGGT AGATGCAAGA CCCGAACGTG
12421 GTGGACCCGG CGCTGCGGGC GGCTCTGCAG AGCCAGCCGT CCGGCCTTAA CTCCTCAGAC
12481 GACTGGCGAC AGGTCATGGA CCGCATCATG TCGCTGACGG CGCGTAACCC GGACGCGTTC
12541 CGGCAGCAGC CGCAGGCCAA CAGGCTCTCC GCCATCCTGG AGGCGGTGGT GCCTGCGCGC
12601 TCGAACCCCA CGCACGAGAA GGTGCTGGCC ATAGTGAACG CGCTGGCCGA GAACAGGGCC
12661 ATCCGCCCCG ACGAGGCCGG GCTGGTGTAC GACGCGCTGC TGCAGCGCGT GGCCCGCTAC
12721 AACAGCGGCA ACGTGCAGAC CAACCTGGAC CGGCTGGTGG GGGACGTGCG CGAGGCGGTG
12781 GCGCAGCGCG AGCGCGCGGA TCGGCAGGGC AACCTGGGCT CCATGGTGGC GCTGAATGCC
12841 TTCCTGAGCA CGCAGCCGGC CAACGTGCCG CGGGGGCAGG AAGACTACAC CAACTTTGTG
12901 AGCGCGCTGC GGCTGATGGT GACCGAGACC CCCAGAGCG AGGTGTACCA GTCGGGCCCC
12961 GACTACTTCT TCCAGACCAG CAGACAGGGC CTGCAGACGG TGAACCTGAG CCAGGCTTTC
13021 AAGAACCTGC GGGGGCTGTG GGGCGTGAAG GCGCCCACCG GCGACCGGGC GACGGTGTCC
13081 AGCCTGCTGA CGCCCAACTC GCGCCTGCTG CTGCTGCTGA TCGCGCCGTT CACGGACAGC
13141 GGCAGCGTGT CCCGGGACAC CTACCTGGGG CACCTGCTGA CCCTGTACCG CGAGGCCATC
13201 GGGCAGGCGC AGGTGGACGA GCACACCTTC CAGGAGATCA CCAGCGTGAG CCGCGCGCTG
13261 GGGCAGGAGG ACACGAGCAG CCTGGAGGCG ACTCTGAACT ACCTGCTGAC CAACCGGCGG
13321 CAGAAGATTC CCTCGCTGCA CAGCCTGACC TCCGAGGAGG AGCGCATCTT GCGCTACGTG
13381 CAGCAGAGCG TGAGCCTGAA CCTGATGCGC GACGGGGTGA CGCCAGCGT GCGCTGGAC
13441 ATGACCGCGC GCAACATGGA ACCGGGCATG TACGCCGCGC ACCGGCCTTA CATCAACCGC
13501 CTGATGGACT ACCTGCATCG CGCGGCGGCC GTGAACCCCG AGTACTTTAC CAACGCCATC
13561 CTGAACCCGC ACTGGCTCCC GCCGCCCGGG TTCTACAGCG GGGGCTTCGA GGTCCCGGAG
13621 GCCAACGATG GCTTCCTGTG GGACGACATG GACGACAGCG TGTTCCTCCC GCGGCCGAG
13681 GCGCTGGCGG AAGCGTCCCT GCTGCGTCCC AAGAAGGAGG AGGAGGAGGC GAGTCGCCG
13741 CGCGGCAGCA GCGGCGTGGC TTCTCTGTCC GAGCTGGGGG CGGCAGCCGC CGCGCGCCCC
13801 GGGTCCCTGG GCGGCAGCCC CTTCCGAGC CTGGTGGGGT CTCTGCACAG CGAGCGCACC
13861 ACCCGCCCTC GGCTGCTGGG CGAGGACGAG TACCTGAATA ACTCCCTGCT GCAGCCGGTG

Fig. 5H

SEQ ID NO:1

13/153

13921 CGGGAGAAAA ACCTGCCCCC CGCCTTCCCC AACAAACGGGA TAGAGAGCCT GGTGGACAAG
13981 ATGAGCAGAT GGAAGACCTA TGCGCAGGAG CACAGGGACG CGCCCGCGCT CCGGCCGCCC
14041 ACGCGGCGCC AGCGCCACGA CCGGCAGCGG GGGCTGGTGT GGGATGACGA GGACTCCGCG
14101 GACGATAGCA GCGTGCTGGA CCTGGGAGGG AGCGGCAACC CGTTCGCGCA CCTGCGCCCC
14161 CGCCTGGGGA GGATGTTTTA AAAAAAAAAA AAGCAAGAAG CATGATGCAA AATTAAATAA
14221 AACTCACCAA GGCCATGGCG ACCGAGCGTT GGTTCCTTGT GTTCCCTTCA GTATGCGGCG
14281 CGCGGCGATG TACCAGGAGG GACCTCCTCC CTCTTACGAG AGCGTGGTGG GCGCGGCGGC
14341 GCGGCGGCCC TCTTCTCCCT TTGCGTCGCA GCTGCTGGAG CCGCCGTACG TGCTTCCGCG
14401 CTACCTGCGG CCTACGGGGG GGAGAAACAG CATCCGTTAC TCGGAGCTGG CGCCCTGTGT
14461 CGACACCACC CGGGTGTACC TGGTGGACAA CAAGTCGGCG GACGTGGCCT CCCTGAACCTA
14521 CCAGAACGAC CACAGCAATT TTTTGACCAC GGTTCATCCAG AACATGACT ACAGCCCGAG
14581 CGAGGCCAGC ACCCAGACCA TCAATCTGGA TGACCGGTCG CACTGGGGCG GCGACCTGAA
14641 AACCATCCTG CACACCAACA TGCCCAACGT GAACGAGTTC ATGTTACCA ATAAGTTCAA
14701 GCGCGGGGTG ATGGTGTGCG GCTCGCACAC CAAGGAAGAC CGGGTGGAGC TGAAGTACGA
14761 GTGGGTGGAG TTCGAGCTGC CAGAGGGCAA CTACTCCGAG ACCATGACCA TTGACCTGAT
14821 GAACAACGCG ATCGTGGAGC ACTATCTGAA AGTGGGCAGG CAAAACGGGG TCCTGGAGAG
14881 CGACATCGGG GTCAAGTTCG ACACCAGGAA CTTCCGCCTG GGGCTGGACC CCGTGACCGG
14941 GCTGGTTATG CCCGGGGTGT ACACCAACGA GGCCTTCCAT CCCGACATCA TCCTGCTGCC
15001 CGGCTGCGGG GTGGACTTCA CTTACAGCCG CCTGAGCAAC CTCCTGGGCA TCCGCAAGCG
15061 GCAGCCCTTC CAGGAGGGCT TCAGGATCAC CTACGAGGAC CTGGAGGGGG GCAACATCCC
15121 CGCGCTCCTC GATGTGGAGG CCTACCAGGA TAGCTTGAAG GAAAATGAGG CGGGACAGGA
15181 GGATACCACC CCCGCCGCT CCGCCGCCGC CGAGCAGGGC GAGGATGCTG CTGACACCGC
15241 GGCCGCGGAC GGGGCAGAGG CCGACCCCGC TATGGTGGTG GAGGCTCCCG AGCAGGAGGA
15301 GGATATGAAT GACAGTGCGG TGCGCGGAGA CACCTTCGTC ACCCGGGGGG AGGAAAAGCA
15361 AGCGGAGGCC GAGGCCGCGG CCGAGGAAAA GCAACTGGCG GCAGCAGCGG CGGCGGCGGC
15421 GTTGGCCGCG GCGGAGGCTG AGTCTGAGGG GACCAAGCCC GCCAAGGAGC CCGTGATTAA
15481 GCCCCTGACC GAAGATAGCA AGAAGCGCAG TTACAACCTG CTCAAGGACA GCACCAACAC
15541 CGCGTACCGC AGCTGGTACC TGGCCTACAA CTACGGCGAC CCGTCGACGG GGGTGCCTC
15601 CTGGACCCTG CTGTGCACGC CGGACGTGAC CTGCGGCTCG GAGCAGGTGT ACTGGTCGCT

Fig. 5I

SEQ ID NO:1

14/153

15661 GCCCGACATG ATGCAAGACC CCGTGACCTT CCGCTCCACG CGGCAGGTCA GCAACTTCCC
15721 GGTGGTGGGC GCCGAGCTGC TGCCCGTGCA CTCCAAGAGC TTCTACAACG ACCAGGCCGT
15781 CTACTCCCAG CTCATCCGCC AGTTCACCTC TCTGACCCAC GTGTTCAATC GCTTTCTCTGA
15841 GAACCAGATT CTGGCGCGCC CGCCCGCCCC CACCATCACC ACCGTCAGTG AAAACGTTCC
15901 TGCTCTCACA GATCACGGGA CGCTACCGCT GCGCAACAGC ATCGGAGGAG TCCAGCGAGT
15961 GACCGTTACT GACGCCAGAC GCCGCACCTG CCCCTACGTT TACAAGGCCT TGGGCATAGT
16021 CTCGCCGCGC GTCCTTTCCA GCCGCACTTT TTGAGCAACA CCACCATCAT GTCCATCCTG
16081 ATCTCACCCA GCAATAACTC CGGCTGGGGA CTGCTGCGCG CGCCAGCAA GATGTTCCGA
16141 GGGGCGAGGA AGCGTTCCGA GCAGCACCCC GTGCGCGTGC GCGGGCACTT CCGCGCCCCC
16201 TGGGGAGCGC ACAAACGCGG CCGCGCGGGG CGCACCACCG TGGACGACGC CATCGACTCG
16261 GTGGTGGAGC AGGCGCGCAA CTACAGGCC GCGGTCTCTA CCGTGGACGC GGCCATCCAG
16321 ACCGTGGTGC GGGGCGCGCG GCGGTACGCC AAGCTGAAGA GCCGCCGGAA GCGCGTGGCC
16381 CGCCGCCACC GCCGCCGACC CGGGGCCGCC GCCAAACGCG CCGCCGCGGC CCTGCTTCGC
16441 CGGGCCAAGC GCACGGGCCG CCGCGCCGCC ATGAGGGCCG CGCGCCGCTT GGCCGCCGGC
16501 ATCACCCCG CCACCATGGC CCCCCGTACC CGAAGACGCG CGGCCGCCGC CGCCGCCGCC
16561 GCCATCAGTG ACATGGCCAG CAGGCGCCGG GGCAACGTGT ACTGGGTGCG CGACTCGGTG
16621 ACCGGCACGC GCGTGCCCGT GCGCTTCCGC CCCCCGCGGA CTTGAGATGA TGTGAAAAA
16681 CAACACTGAG TCTCCTGCTG TTGTGTGTAT CCCAGCGGCG GCGGCGCGCG CAGCGTCATG
16741 TCCAAGCGCA AAATCAAAGA AGAGATGCTC CAGGTCGTG CGCCGGAGAT CTATGGGCCC
16801 CCGAAGAAGG AAGAGCAGGA TTCGAAGCCC CGCAAGATAA AGCGGGTCAA AAAGAAAAAG
16861 AAAGATGATG ACGATGCCGA TGGGAGGTG GAGTTCCTGC GCGCCACGGC GCCCAGGCGC
16921 CCGGTGCAGT GGAAGGGCCG GCGCGTAAAG CGCGTCCTGC GCCCCGGCAC CGCGGTGGTC
16981 TTCACGCCCC GCGAGCGCTC CACCCGGA CTCAAGCGCG TCTATGACGA GGTGTACGGC
17041 GACGAAGACC TGCTGGAGCA GGCCAACGAG CGCTTCGGAG AGTTTGCTTA CGGGAAGCGT
17101 CAGCGGGCGC TGGGGAAGGA GGACCTGCTG GCGCTGCCGC TGGACCAGGG CAACCCACC
17161 CCCAGTCTGA AGCCCGTGAC CCTGCAGCAG GTGCTGCCGA GCAGCGCACC CTCCGAGGCG
17221 AAGCGGGGTC TGAAGCGCGA GGGCGGCGAC CTGGCGCCCA CCGTGCAGCT CATGGTGCCC
17281 AAGCGGCAGA GGCTGGAGGA TGTGCTGGAG AAAATGAAAG TAGACCCCGG TCTGCAGCCG
17341 GACATCAGGG TCCGTCCCAT CAAGCAGGTG GCGCCGGGCC TCGGCGTGCA GACCGTGGAC

Fig. 5J

SEQ ID NO:1

15/153

17401 GTGGTCATCC CCACCGGCAA CTCCCCCGCC GCCACCACCA CTACCGCTGC CTCCACGGAC
17461 ATGGAGACAC AGACCGATCC CGCCGCAGCC GCAGCCGCCG CCGCAGCCGC GACCTCCTCG
17521 GCGGAGGTGC AGACGGACCC CTGGCTGCCG CCGGCGATGT CAGCTCCCCG CGCGCGCCGC
17581 GGACGCAGAA AGTACGGCGC CGCCAACGCG CTCCTGCCCG AGTACGCCTT GCATCCTTCC
17641 ATCGCGCCCA CCCCCGGCTA CCGAGGCTAT ACCTACCGCC CGCGAAGAGC CAAGGGTTCC
17701 ACCCGCCGTC CCGCCGACG CGCCGCCGCC ACCACCCGCC GCGCGCGCCG CAGACGCCAG
17761 CCCGCACTGG CTCCAGTCTC CGTGAGGAGA GTGGCGCGCG ACGGACACAC CCTGGTGCTG
17821 CCCAGGGCGC GCTACCACCC CAGCATCGTT TAAAAGCCTG TTGTGGTTCT TGCAGATATG
17881 GCCCTCACTT GCCGCCTCCG TTTCCCGGTG CCGGGATACC GAGGAGGAAG ATCGCGCCGC
17941 AGGAGGGGTC TGGCCGGCCG CGGCCTGAGC GGAGGCAGCC GCCGCGCGCA CCGGCGGCGA
18001 CGCGCCACCA GCCGACGCAT GCGCGGCGGG GTGCTGCCCC TGTTAATCCC CCTGATCGCC
18061 GCGGCGATCG GCGCCGTGCC CGGGATCGCC TCCGTGGCCT TGCAAGCGTC CCAGAGGCAT
18121 TGACAGACTT GCAAACCTGC AAATATGGAA AAAAAAAAAA AACCCCAATA AAAAGTCTAG
18181 ACTCTCACGC TCGCTTGGTC CTGTGACTAT TTTGTAGAAT GGAAGACATC AACTTTGCGT
18241 CGCTGGCCCC GCGTCACGGC TCGCGCCCGT TCCTGGGACA CTGGAACGAT ATCGGCACCA
18301 GCAACATGAG CGGTGGCGCC TTCAGTTGGG GCTCTCTGTG GAGCGGCATT AAAAGTATCG
18361 GGTCTGCCGT TAAAAATTAC GGCTCCCGGG CCTGGAACAG CAGCACGGGC CAGATGTTGA
18421 GAGACAAGTT GAAAGAGCAG AACTTCCAGC AGAAGGTGGT GGAGGGCCTG GCCTCCGGCA
18481 TCAACGGGGT GGTGGACCTG GCCAACCAGG CCGTGCAGAA TAAAATCAAC AGCAGACTGG
18541 ACCCCCGGCC GCCGGTGGAG GAGGTGCCGC CGGCGCTGGA GACGGTGTCC CCCGATGGGC
18601 GTGGCGAGAA GCGCCCGCGG CCCGATAGGG AAGAGACCAC TCTGGTCACG CAGACCGATG
18661 AGCCGCCCCC GTATGAGGAG GCCCTAAAGC AAGGTCTGCC CACCACGCGG CCCATCGCGC
18721 CCATGGCCAC CGGGGTGGTG GGCCGCCACA CCCCCGCCAC GCTGGACTTG CCTCCGCCCC
18781 CCGATGTGCC GCAGCAGCAG AAGGCGGCAC AGCCGGGCCC GCGCGGACC GCCTCCCGTT
18841 CCTCCGCCGG TCCTCTGCGC CGCGCGGCCA GCGGCCCCCG CGGGGGGGTC GCGAGGCACG
18901 GCAACTGGCA GAGCACGCTG AACAGCATCG TGGGTCTGGG GTGCGGTCC GTGAAGCGCC
18961 GCCGATGCTA CTGAATAGCT TAGCTAACGT GTTGATGTG TGATGCGCC CTATGTCGCC
19021 GCCAGAGGAG CTGCTGAGTC GCCGCCGTTT GCGCGCCAC CACCACCGCC ACTCCGCCCC
19081 TCAAGATGGC GACCCCATCG ATGATGCCGC AGTGGTCGTA CATGCACATC TCGGGCCAGG

Fig. 5K

SEQ ID NO:1

16/153

19141 ACGCCTCGGA GTACCTGAGC CCCGGGCTGG TGCAGTTCGC CCGCGCCACC GAGAGCTACT
19201 TCAGCCTGAG TAACAAGTTT AGGAACCCCA CGGTGGCGCC CACGCACGAT GTGACCACCG
19261 ACCGGTCTCA GCGCCTGACG CTGCGGTTCA TTCCCGTGGA CCGCGAGGAC ACCGCGTACT
19321 CGTACAAGGC GCGGTTCAAC CTGGCCGTGG GCGACAACCG CGTGCTGGAC ATGGCCTCCA
19381 CCTACTTTGA CATCCGCGGG GTGCTGGACC GGGGTCCCAC TTTCAGCCC TACTCTGGCA
19441 CCGCCTACAA CTCCCTGGCC CCCAAGGGCG CTCCCAACTC CTGCGAGTGG GAGCAAGAGG
19501 AAACCTCAGGC AGTTGAAGAA GCAGCAGAAG AGGAAGAAGA AGATGCTGAC GGTCAAGCTG
19561 AGGAAGAGCA AGCAGCTACC AAAAAGACTC ATGTATATGC TCAGGCTCCC CTTTCTGGCG
19621 AAAAAATTAG TAAAGATGGT CTGCAAATAG GAACGGACGC TACAGCTACA GAACAAAAAC
19681 CTATTTATGC AGACCTTACA TTCCAGCCCG AACCCCAAAT CGGGGAGTCC CAGTGGAATG
19741 AGGCAGATGC TACAGTCGCC GCGGGTAGAG TGCTAAAGAA ATCTACTCCC ATGAAACCAT
19801 GCTATGGTTC CTATGCAAGA CCCACAAATG CTAATGGAGG TCAGGGTGTA CTAACGGCAA
19861 ATGCCCAGGG ACAGCTAGAA TCTCAGGTTG AAATGCAATT CTTTTCAACT TCTGAAAACG
19921 CCCGTAACGA GGCTAACAAC ATTCAGCCCA AATTGGTGCT GTATAGTGAG GATGTGCACA
19981 TGGAGACCCC GGATACGCAC CTTTCTTACA AGCCCGCAA AAGCGATGAC AATTCAAAAA
20041 TCATGCTGGG TCAGCAGTCC ATGCCCAACA GACCTAATTA CATCGGCTTC AGAGACAACCT
20101 TTATCGGCCT CATGTATTAC AATAGCACTG GCAACATGGG AGTGCTTGCA GGTCAAGCCT
20161 CTCAGTTGAA TGCAAGTGGT GACTTGCAAG ACAGAAACAC AGAACTGTCC TACCAGCTCT
20221 TGCTTGATTC CATGGGTGAC AGAACCAGAT ACTTTTCCAT GTGGAATCAG GCAGTGGACA
20281 GTTATGACCC AGATGTTAGA ATTATTGAAA ATCATGGAAC TGAAGACGAG CTCCCCAAT
20341 ATTGTTTCCC TCTGGGTGGC ATAGGGGTAA CTGACACTTA CCAGGCTGTT AAAACCAACA
20401 ATGGCAATAA CGGGGGCCAG GTGACTTGGA CAAAAGATGA AACTTTTGCA GATCGCAATG
20461 AAATAGGGGT GGGAAACAAT TTCGCTATGG AGATCAACCT CAGTGCCAAC CTGTGGAGAA
20521 ACTTCCTGTA CTCCAACGTG GCGCTGTACC TACCAGACAA GCTTAAGTAC AACCCCTCCA
20581 ATGTGGACAT CTCTGACAAC CCCAACACCT ACGATTACAT GAACAAGCGA GTGGTGGCCC
20641 CGGGGCTGGT GGACTGCTAC ATCAACCTGG GCGCGCGCTG GTCGCTGGAC TACATGGACA
20701 ACGTCAACCC CTTCAACCAC CACCGCAATG CGGGCCTGCG CTACCGCTCC ATGCTCCTGG
20761 GCAACGGGCG CTACGTGCCC TTCCACATCC AGGTGCCCCA GAAGTTCTTT GCCATCAAGA
20821 ACCTCCTCCT CCTGCCGGGC TCCTACACCT ACGAGTGGA CTTAGGAAG GATGTCAACA

Fig. 5L

SEQ ID NO:1

17/153

20881 TGGTCCTCCA GAGCTCTCTG GGTAACGATC TCAGGGTGGA CGGGGCCAGC ATCAAGTTCG
20941 AGAGCATCTG CCTCTACGCC ACCTTCTTCC CCATGGCCCA CAACACGGCC TCCACGCTCG
21001 AGGCCATGCT CAGGAACGAC ACCAACGACC AGTCCTTCAA TGACTACCTT TCCGCCGCCA
21061 ACATGCTCTA CCCCATACCC GCCAACGCCA CCAACGTCCC CATCTCCATC CCCTCGCGCA
21121 ACTGGGCGGC CTTCCGCGGC TGGGCCTTCA CCCGCCTCAA GACCAAGGAG ACCCCCTCCC
21181 TGGGCTCGGG ATTCGACCCC TACTACACCT ACTCGGGCTC TATTCCTTAC CTGGACGGCA
21241 CCTTCTACCT CAACCACACT TTCAAGAAGG TCTCGGTCAC CTTGACTCC TCGGTCAGCT
21301 GGCCGGGCAA CGACCGTCTG CTCACCCCA ACGAGTTCGA GATCAAGCGC TCGGTCGACG
21361 GGGGAAGGCTA CAACGTGGCC CAGTGCAACA TGACCAAGGA CTGGTTCTTG GTCCAGATGC
21421 TGGCCAACTA CAACATCGGC TACCAGGGCT TCTACATCCC AGAGAGCTAC AAGGACAGGA
21481 TGTACTCCTT CTTCAGGAAC TTCCAGCCCA TGAGCCGGCA GGTGGTGGAC CAGACCAAGT
21541 ACAAGGACTA CCAGGAGGTG GGCATCATCC ACCAGCACA CAACTCGGGC TTCGTGGGCT
21601 ACCTCGCCCC CACCATGCGC GAGGGACAGG CCTACCCCGC CAACTTCCCC TACCCGCTCA
21661 TAGGCAAGAC CGCGGTCGAC AGCATCACCC AGAAAAAGTT CCTCTGCGAC CGCACCTCT
21721 GGCGCATCCC CTCTCCAGC AACTTCATGT CCATGGGTGC GCTCTCGGAC CTGGGCCAGA
21781 ACTTGCTCTA CGCCAACTCC GCCACGCCC TCGACATGAC CTTGAGGTC GACCCCATGG
21841 ACGAGCCAC CCTTCTCTAT GTTCTGTTG AAGTCTTTGA CGTGGTCCGG GTCCACCAGC
21901 CGCACCGCGG CGTCATCGAG ACCGTGTACC TGCGTACGCC CTTCTCGGCC GGCAACGCCA
21961 CCACCTAAAG AAGCAAGCCG CAGTCATCGC CGCCTGCATG CCGTCGGGTT CCACCGAGCA
22021 AGAGCTCAGG GCCATCGTCA GAGACCTGGG ATGCGGGCCC TATTTTTTGG GCACCTTCGA
22081 CAAGCGCTTC CCTGGCTTTG TCTCCCCACA CAAGCTGGCC TGCGCCATCG TCAACACGGC
22141 CGGCCGCGAG ACCGGGGGCG TGCACTGGCT GGCCTTTGCC TGGAACCCGC GCTCCAAAAC
22201 ATGCTTCCTC TTTGACCCCT TCGGCTTTTC GGACCAGCGG CTCAAGCAA TCTACGAGTT
22261 CGAGTACGAG GGCTTGCTGC GTCGACGCG CATCGCCTCC TCGCCCGACC GCTGCGTCAC
22321 CCTCGAAAAG TCCACCCAGA CCGTGCAGGG GCCCGACTCG GCCGCTGCG GTCTCTTCTG
22381 CTGCATGTTT CTGCACGCCT TTGTGCACTG GCCTCAGAGT CCCATGGACC GCAACCCAC
22441 CATGAACCTG CTGACGGGGG TGCCAACTC CATGCTCAA AGCCCCAGG TCGAGCCCAC
22501 CCTGCGCCGC AACCAGGAGC AGCTCTACAG CTTCTGGAG CGCCACTCGC CCTACTTCCG
22561 CCGCCACAGC GCACAGATCA GGAGGGCCAC CTCCTTCTGC CACTTGCAAG AGATGCAAGA

Fig. 5M

SEQ ID NO:1

18/153

22621 AGGGTAATAA CGATGTACAC ACTTTTTTCT CAATAAATGG CATTTTTTTT TTATTTATAC
22681 AAGCTCTCTG GGGTATTCAT TTCCCACCAC CACCACCCGC CGTTGTCGCC ATCTGGCTCT
22741 ATTTAGAAAT CGAAAGGGTT CTGCCGGGAG TCGCCGTGCG CCACGGGCAG GGACACGTTG
22801 CGATACTGGT AGCGGGTGCC CCACTTGAAC TCGGGCACCA CCAGGCGAGG CAGCTCGGGG
22861 AAGTTTTTCG TCCACAGGCT GCGGGTCAGC ACCAGCGCGT TCATCAGGTC GGGCGCCGAG
22921 ATCTTGAAGT CGCAGTTGGG GCCGCCGCC TCGCGCGCG AGTTGCGGTA CACCGGGTTG
22981 CAGCACTGGA ACACCAACAG CGCCGGGTGC TTCACGCTGG CCAGCACGCT GCGGTCGGAG
23041 ATCAGCTCGG CGTCCAGGTC CTCCGCGTTG CTCAGCGCGA ACGGGGTCAT CTTGGGCACT
23101 TGCCGCCCCA GGAAGGGCGC GTGCCCCGGT TTCGAGTTGC AGTCGCAGCG CAGCGGGATC
23161 AGCAGGTGCC CGTGCCCGGA CTCGGCGTTG GGGTACAGCG CGCGCATGAA GGCCTGCATC
23221 TGGCGGAAGG CCATCTGGGC CTTGGCGCCC TCCGAGAAGA ACATGCCGCA GGACTTGCCC
23281 GAGAACTGGT TTGCGGGGCA GCTGGCGTCG TGCAGGCAGC AGCGCGCGTC GGTGTTGGCG
23341 ATCTGCACCA CGTTGCGCCC CCACCGGTTT TTCACGATCT TGGCCTTGGA CGATTGCTCC
23401 TTCAGCGCGC GCTGCCCCTT CTCGCTGGTC ACATCCATCT CGATCACATG TTCCTTGTTT
23461 ACCATGCTGC TGCCGTGCAG ACACTTCAGC TCGCCCTCCG TCTCGGTGCA GCGGTGCTGC
23521 CACAGCGCGC AGCCCGTGGG CTCGAAAGAC TTGTAGGTCA CCTCCGCGAA GGACTGCAGG
23581 TACCCCTGCA AAAAGCGGCC CATCATGGTC ACGAAGGTCT TGTGCTGCT GAAGGTCAGC
23641 TGCAGCCCGC GGTGCTCCTC GTTCAGCCAG GTCTTGACA CGGCCGCCAG CGCCTCCACC
23701 TGGTCGGGCA GCATCTTGAA GTTCACCTT AGCTCATTCT CCACGTGGTA CTTGTCCATC
23761 AGCGTGCGCG CCGCCTCCAT GCCCTTCTCC CAGGCCGACA CCAGCGGCAG GCTCACGGGG
23821 TTCTTCACCA TCACCGTGGC CGCCGCCTCC GCCGCGCTTT CGCTTTCCGC CCCGCTGTTT
23881 TCTTCTCTT CCTCCTCTT CTCGCCGCC CCCACTCGCA GCGCCCGCAC CACGGGGTCG
23941 TCTTCTGCA GCGCTGCAC CTTGCGCTT CCGTTGCGCC CCTGCTTGAT GCGCACGGGC
24001 GGGTTGCTGA AGCCACCAT CACCAGCGC GCCTCTTCTT GCTCGTCTC GCTGTCCAGA
24061 ATGACCTCCG GGGAGGGGG GTTGGTCATC CTCAGTACCG AGGCACGCTT CTTTTTCTT
24121 CTGGGGGCGT TCGCCAGCTC CGCGGCTGCG GCCGCTGCC AGGTGGAAGG CCGAGGGCTG
24181 GCGGTGCGCG GCACCAGCG GTCTTGCGAG CCGTCTCGT CCTCCTCGGA CTCGAGACGG
24241 AGGCGGGCCC GCTTCTTCGG GGGCGCGCG GCGGCGGAG GCGGCGGCG CGACGGAGAC
24301 GGGGACGAGA CATCGTCCAG GGTGGGTGGA CGGCGGGCCG CGCCGCGTCC GCGCTCGGGG

Fig. 5N

SEQ ID NO:1

19/153

24361 GTGGTTTCGC GCTGGTCCTC TTCCCGACTG GCCATCTCCC ACTGCTCCTT CTCCTATAGG
24421 CAGAAAGAGA TCATGGAGTC TCTCATGCGA GTCGAGAAGG AGGAGGACAG CCTAACCGCC
24481 CCCTCTGAGC CCTCCACCAC CGCCGCCACC ACCGCCAATG CCGCCGCGGA CGACGCGCCC
24541 ACCGAGACCA CCGCCAGTAC CACCCTCCCC AGCGACGCAC CCCCCTCGA GAATGAAGTG
24601 CTGATCGAGC AGGACCCGGG TTTTGTGAGC GGAGAGGAGG ATGAGGTGGA TGAGAAGGAG
24661 AAGGAGGAGG TCGCCGCCCTC AGTGCCAAAA GAGGATAAAA AGCAAGACCA GGACGACGCA
24721 GATAAGGATG AGACAGCAGT CGGGCGGGGG AACGGAAGCC ATGATGCTGA TGACGGCTAC
24781 CTAGACGTGG GAGACGACGT GCTGCTTAAG CACCTGCACC GCCAGTGCGT CATCGTCTGC
24841 GACGCGCTGC AGGAGCGCTG CGAAGTGCCC CTGGACGTGG CGGAGGTCAG CCGCGCCTAC
24901 GAGCGGCACC TCTTCGCGCC GCACGTGCCC CCAAGCGCC GGGAGAACGG CACCTGCGAG
24961 CCAACCCGC GTCTCAACTT CTACCCGGTC TTCGCGGTAC CCGAGGTGCT GGCCACCTAC
25021 CACATCTTCT TCCAAACTG CAAGATCCCC CTCTCCTGCC GCGCTAACCG CACCCGCGCC
25081 GACAAAACCC TGACCCTGCG GCAGGGCGCC CACATACCTG ATATTGCCTC TCTGGAGGAA
25141 GTGCCCAAGA TCTTCGAGGG TCTCGGTCGC GACGAGAAAC GGGCGGCGAA CGCTCTGCAC
25201 GGAGACAGCG AAAACGAGAG TCACTCGGGG GTGCTGGTGG AGCTCGAGGG CGACAACGCG
25261 CGCCTGGCCG TACTCAAGCG CAGCATAGAG GTCACCCACT TTGCCTACCC GGCGCTCAAC
25321 CTGCCCCCA AGGTCATGAG TGTGGTCATG GGCGAGCTCA TCATGCGCCG CGCTCAGCCC
25381 CTGGCCGCGG ATGCAAACTT GCAAGAGTCC TCCGAGGAAG GCCTGCCCCG GGTCAGCGAC
25441 GAGCAGCTAG CGCGCTGGCT GGAGACCCGC GACCCGCGC AGCTGGAGGA GCGGCGCAAG
25501 CTCATGATGG CCGCGGTGCT GGTACCCGTG GAGCTCGAGT GTCTGCAGCG CTCTTCGCG
25561 GACCCCGAGA TGCAGCGCAA GCTCGAGGAG ACCCTGCACT ACACCTTCCG CCAGGGCTAC
25621 GTGCGCCAGG CCTGCAAGAT CTCCAACGTG GAGCTCTGCA ACCTGGTCTC CTACCTGGGC
25681 ATCCTGCACG AGAACCGCCT CGGGCAGAAC GTCCTGCACT CCACCCTCAA AGGGGAGGCG
25741 CGCCGCGACT ACATCCGCGA CTGCGCCTAC CTCTTCTCT GCTACACCTG GCAGACGGCC
25801 ATGGGGGTCT GGCAGCAGTG CCTGGAGGAG CGCAACCTCA AGGAGCTGGA AAAGCTACTC
25861 AAGCGACCC TCAGGGACCT CTGGACGGGC TTCAACGAGC GCTCGGTGGC CGCCGCGCTG
25921 GCGGACATCA TCTTCCCCGA GCGCCTGCTC AAGACCCTGC AGCAGGGCCT GCCCGACTTC
25981 ACCAGCCAGA GCATGCTGCA GAACTTTAGG ACTTTCATCC TGGAGCGCTC GGGCATCTG
26041 CCTGCCACTT GCTGCGCGCT GCCCAGCGAC TTCGTGCCCA TCAAGTACAG GGAGTGCCCG

Fig. 50

SEQ ID NO:1

20/153

26101 CCGCCGCTCT GGGGCCACTG CTACCTCTTC CAGCTGGCCA ACTACCTCGC CTACCACTCG
26161 GACCTCATGG AAGACGTGAG CGGCGAGGGC CTGCTCGAGT GCCACTGCCG CTGCAACCTC
26221 TGCACGCCCC ACCGCTCTCT AGTCTGCAAC CCGCAGCTGC TCAGCGAGAG TCAGATTATC
26281 GGTACCTTCG AGCTGCAGGG TCCCTCGCCT GACGAGAAGT CCGCGGCTCC GGGGCTGAAA
26341 CTCACTCCGG GGCTGTGGAC TTCCGCCTAC CTACGCAAAT TTGTACCTGA GGACTACCAC
26401 GCCCAGGAGA TCAGGTTCTA CGAAGACCAA TCCCCCCCCG CCAAGGCGGA GCTCACCGCC
26461 TGCCTCATCA CCCAGGGGCA CATCTGGGC CAATTGCAAG CCATCAACAA AGCCCCCGCA
26521 GAGTTCTTGC TGAAAAGGG TCGGGGGGTG TACCTGGACC CCCAGTCCGG CGAGGAGCTA
26581 AACCCGCTAC CCCC GCCGCC GCCCCAGCAG CGGGACCTTG CTTCCCAGGA TGGCACCCAG
26641 AAAGAAGCAG CAGCCGCCGC CGCCGAGCC ATACATGCTT CTGGAGGAAG AGGAGGAGGA
26701 CTGGGACAGT CAGGCAGAGG AGGTTTCGGA CGAGGAGCAG GAGGAGATGA TGGAAAGACTG
26761 GGAGGAGGAC AGCAGCCTAG ACGAGGAAGC TTCAGAGGCC GAAGAGGTGG CAGACGCAAC
26821 ACCATCACCC TCGGTCGCAG CCCCCTCGCC GGGGCCCCCTG AAATCCTCCG AACCCAGCAC
26881 CAGCGCTATA ACCTCCGCTC CTCGGCGGCC GCGGCCACCC GCCCGCAGAC CCAACCGTAG
26941 ATGGGACACC ACAGGAACCG GGGTCGGTAA GTCCAAGTGC CCGCCGCCGC CACCGCAGCA
27001 GCAGCAGCAG CGCCAGGGCT ACCGCTCGTG GCGCGGGCAC AAGAACGCCA TAGTCGCCTG
27061 CTTGCAAGAC TCGGGGGGCA ACATCTCTTT CGCCCGGCGC TTCCTGCTAT TCCACCACGG
27121 GGTGCGCTTT CCCC GCAATG TCCTGCATTA CTACCGTCAT CTCTACAGCC CCTACTGCAG
27181 CGGCGACCCA GAGGCGGCAG CGGCAGCCAC AGCGGCGACC ACCACCTAGG AAGATATCCT
27241 CCGCGGGCAA GACAGCGGCA GCAGCGGCCA GGAGACCCGC GGCAGCAGCG GCGGGAGCGG
27301 TGGGCGCACT GCGCCTCTCG CCCAACGAAC CCCTCTCGAC CCGGGAGCTC AGACACAGGA
27361 TCTTCCCCAC TTTGTATGCC ATCTTCCAAC AGAGCAGAGG CCAGGAGCAG GAGCTGAAAA
27421 TAAAAAACAG ATCTCTGCGC TCCCTCACCC GCAGCTGTCT GTATCACAAA AGCGAAGATC
27481 AGCTTCGGCG CACGCTGGAG GACGCGGAGG CACTCTTCAG CAAATACTGC GCGCTCACTC
27541 TTAAAGACTA GCTCCGCGCC CTTCTCGAAT TTAGGCGGGA GAAACTACG TCATCGCCGG
27601 CCGCCGCCCA GCCCGCCCAG CCGAGATGAG CAAAGAGATT CCCACGCCAT ACATGTGGAG
27661 CTACCAGCCG CAGATGGGAC TCGCGGCGGG AGCGGCCAG GACTACTCCA CCCGCATGAA
27721 CTACATGAGC GCGGGACCCC ACATGATCTC ACAGGTCAAC GGGATCCGCG CCCAGCGAAA
27781 CCAAATACTG CTGGAACAGG CGGCCATCAC CGCCACGCC CGCCATAATC TCAACCCCCG

Fig. 5P

SEQ ID NO:1

21/153

27841 AAATTGGCCC GCCGCCCTCG TGTACCAGGA AACCCCCTCC GCCACCACCG TACTACTTCC
27901 GCGTGACGCC CAGGCCGAAG TCCAGATGAC TAACTCAGGG GCGCAGCTCG CGGGCGGCTT
27961 TCGTCACGGG GCGCGGCCGC TCCGACCAGG TATAAGACAC CTGATGATCA GAGGCCGAGG
28021 TATCCAGCTC AACGACGAGT CGGTGAGCTC TTCGCTCGGT CTCCGTCCGG ACGGAACTTT
28081 CCAGCTCGCC GGATCCGGCC GCTCTTCGTT CACGCCCCGC CAGGCGTACC TGA CTCTGCA
28141 GACCTCGTCC TCGGAGCCCC GCTCCGGAGG CATCGGAACC CTCCAGTTCG TGGAGGAGTT
28201 CGTGCCCTCG GTCTACTTCA ACCCCTTCTC GGGACCTCCC GGACGCTACC CCGACCAGTT
28261 CATTCGGAAC TTTGACGCGG TGAAGGACTC GGCGGACGGC TACGACTGAA TGTCAGGTGC
28321 CGAGGCAGAG CAGCTTCGCC TGAGACACCT CGAGCACTGC CGCCGCCACA AGTGCTTCGC
28381 CCGCGGTTCC GGTGAGTTCT GCTACTTTCA GCTACCCGAG GAGCATACCG AGGGGCCGGC
28441 GCACGGCGTC CGCCTGACCA CCCAGGGCGA GGTACCTGT TCCCTCATCC GGGAGTTCAC
28501 CCTCCGTCCC CTGCTAGTGG AGCGGGAGCG GGGTCCCTGT GTCTAACTA TCGCCTGCAA
28561 CTGCCCTAAC CCTGGATTAC ATCAAGATCT TTGCTGTCAT CTCTGTGCTG AGTTTAATAA
28621 ACGCTGAGAT CAGAATCTAC TGGGGCTCCT GTCGCCATCC TGTGAACGCC ACCGTCTTCA
28681 CCCACCCGA CCAGGCCAG GCGAACCTCA CCTGCGGTCT GCATCGGAGG GCCAAGAAGT
28741 ACCTCACCTG GTACTTCAAC GGCACCCCTT TTGTGGTTTA CAACAGCTTC GACGGGGACG
28801 GAGTCTCCCT GAAAGACCAG CTCTCCGGTC TCAGCTACTC CATCCACAAG AACACCACC
28861 TCCAACCTT CCCTCCCTAC CTGCCGGGAA CCTACGAGTG CGTCACCGGC CGCTGCACCC
28921 ACCTACCCG CCTGATCGTA AACCAGAGCT TTCCGGGAAC AGATAACTCC CTCTTCCCCA
28981 GAACAGGAGG TGAGCTCAGG AACTCCCCG GGGACCAGG CGGAGACGTA CCTTCGACCC
29041 TTGTGGGGTT AGGATTTTTT ATTACCGGGT TGCTGGCTCT TTTAATCAA GCTTCCTTGA
29101 GATTGTCTT TTCCTTCTAC GTGTATGAAC ACCTCAGCCT CCAATAACTC TACCCCTTCT
29161 TCGGAATCAG GTGACTTCTC TGAAATCGGG CTTGGTGTGC TGCTTACTCT GTTGATTTTT
29221 TTCCTTATCA TACTCAGCCT TCTGTGCCTC AGGCTCGCCG CTGCTGCGC ACACATCTAT
29281 ATCTACTGCT GGTTGCTCAA GTGCAGGGGT CGCCACCCAA GATGAACAGG TACATGGTCC
29341 TATCGATCCT AGGCCTGCTG GCCCTGGCGG CCTGCAGCGC CGCCAAAAA GAGATTACCT
29401 TTGAGGAGCC CGCTTGCAAT GTAACCTTCA AGCCCGAGGG TGACCAATGC ACCACCCTCG
29461 TCAAATGCGT TACCAATCAT GAGAGGCTGC GCATCGACTA CAAAAACAAA ACTGGCCAGT
29521 TTGCGGTCTA TAGTGTGTTT ACGCCCGGAG ACCCCTCTAA CTACTCTGTC ACCGTCTTCC

Fig. 5Q

SEQ ID NO:1

22/153

29581 AGGGCGGACA GTCTAAGATA TTCAATTACA CTTTCCCTTT TTATGAGTTA TGCGATGCGG
29641 TCATGTACAT GTCAAAACAG TACAACCTGT GGCCTCCCTC TCCCCAGGCG TGTGTGGAAA
29701 ATACTGGGTC TTACTGCTGT ATGGCTTTGG CAATCACTAC GCTCGCTCTA ATCTGCACGG
29761 TGCTATACAT AAAATTCAGG CAGAGGCGAA TCTTTATCGA TGAAAAGAAA ATGCCTTGAT
29821 CGCTAACACC GGCTTTCTAT CTGCAGAATG AATGCAATCA CCTCCCTACT AATCACCACC
29881 ACCCTCCTTG CGATTGCCCA TGGGTTGACA CGAATCGAAG TGCCAGTGGG GTCCAATGTC
29941 ACCATGGTGG GCCCCGCCGG CAATTCCACC CTCATGTGGG AAAAATTTGT CCGCAATCAA
30001 TGGGTTTCATT TCTGCTCTAA CCGAATCAGT ATCAAGCCCA GAGCCATCTG CGATGGGCAA
30061 AATCTAACTC TGATCAATGT GCAAATGATG GATGCTGGGT ACTATTACGG GCAGCGGGGA
30121 GAAATCATT AATTACTGGCG ACCCCACAAG GACTACATGC TGCATGTAGT CGAGGCACTT
30181 CCCACTACCA CCCCCACTAC CACCTCTCCC ACCACCACTA CCACCACTAC TACTACTACT
30241 ACTACCACTA CCGCTGCCCCG CCATACCCGC AAAAGCACCA TGATTAGCAC AAAGCCCCCT
30301 CGTGCTCACT CCCACGCCGG CGGGCCCATC GGTGCGACCT CAGAAACCAC CGAGCTTTGC
30361 TTCTGCCAAT GACTAACGC CAGCGCTCAT GAACTGTTCG ACCTGGAGAA TGAGGATGCC
30421 CAGCAGAGCT CCGCTTGCCCT GACCCAGGAG GCTGTGGAGC CCGTTGCCCT GAAGCAGATC
30481 GGTGATTCAA TAATTGACTC TTCTTCTTTT GCCACTCCCG AATACCCTCC CGATTCTACT
30541 TTCCACATCA CGGGTACCAA AGACCCTAAC CTCTCTTTCT ACCTGATGCT GCTGCTCTGT
30601 ATCTCTGTGG TCTCTTCCGC GCTGATGTTA CTGGGGATGT TCTGCTGCCT GATCTGCCGC
30661 AGAAAGAGAA AAGCTCGCTC TCAGGGCCAA CCACTGATGC CCTTCCCCTA CCCCCGGAT
30721 TTTGCAGATA ACAAGATATG AGCTCGCTGC TGACACTAAC CGCTTTACTA GCCTGCGCTC
30781 TAACCCTTGT CGCTTGCGAC TCGAGATTCC ACAATGTCAC AGCTGTGGCA GGAGAAAATG
30841 TTACTTTCAA CTCCACGGCC GATACCCAGT GGTGCTGGAG TGGCTCAGGT AGCTACTTAA
30901 CTATCTGCAA TAGTCCACT TCCCCAGCA TATCCCCAAC CAAGTACCAA TGCAATGCCA
30961 GCCTGTTTAC CCTCATCAAC GCTTCCACCC TGGACAATGG ACTCTATGTA GGCTATGTAC
31021 CCTTTGGTGG GCAAGGAAAG ACCCACGCTT ACAACCTGGA AGTTCGCCAG CCCAGAACCA
31081 CTACCCAAGC TTCTCCACC ACCACCACCA CCACCACCAC CACCATCACC AGCAGCAGCA
31141 GCAGCCACAG CAGCAGCAGC AGATTATTGA CTTTGGTTTT GGCCAGCTCA TCTGCCGCTA
31201 CCCAGGCCAT CTACAGCTCT GTGCCCCGAA CCACTCAGAT CCACCGCCCA GAAACGACCA
31261 CCGCCACCAC CCTACACACC TCCAGCGATC AGATGCCGAC CAACATCACC CCCTTGGCTC

Fig. 5R

SEQ ID NO:1

23/153

31321 TTCAAATGGG ACTTACAAGC CCCACTCCAA AACCAGTGGA TGCGGCCGAG GTCTCCGCCC
31381 TCGTCAATGA CTGGGCGGGG CTGGGAATGT GGTGGTTCGC CATAGGCATG ATGGCGCTCT
31441 GCCTGCTTCT GCTCTGGCTC ATCTGCTGCC TCCACCGCAG GCGAGCCAGA CCCCCATCT
31501 ATAGACCCAT CATTGTCCTG AACCCCGATA ATGATGGGAT CCATAGATTG GATGGCCTGA
31561 AAAACCTACT TTTTCTTTT ACAGTATGAT AAATTGAGAC ATGCCTCGCA TTTTCTTGTA
31621 CATGTTCCCT CTCCCACCTT TTCTGGGGTG TTCTACGCTG GCCGCTGTGT CTCACCTGGA
31681 GGTAGACTGC CTCTCACCTT TCACTGTCTA CCTGCTTTAC GGATTGGTCA CCCTCACTCT
31741 CATCTGCAGC CTAATCACAG TAATCATCGC CTTTCATCCAG TGCATTGATT ACATCTGTGT
31801 GCGCCTCGCA TACTTCAGAC ACCACCCGCA GTACCGAGAC AGGAACATTG CCCAACTTCT
31861 AAGACTGCTC TAATCATGCA TAAGACTGTG ATCTGCCTTC TGATCCTCTG CATCCTGCCC
31921 ACCCTCACCT CCTGCCAGTA CACCACAAA TCTCCGCGCA AAAGACATGC CTCCTGCCGC
31981 TTCACCCAAC TGTGGAATAT ACCCAAATGC TACAACGAAA AGAGCGAGCT CTCCGAAGCT
32041 TGGCTGTATG GGGTCATCTG TGTCTTAGTT TTCTGCAGCA CTGTCTTTGC CCTCATGATC
32101 TACCCCTACT TTGATTTGGG ATGGAACGCG ATCGATGCCA TGAATTACCC CACCTTTCCC
32161 GCACCCGAGA TAATTCCACT GCGACAAGTT GTACCCGTTG TCGTTAATCA ACGCCCCCA
32221 TCCCCCTACGC CCACTGAAAT CAGCTACTTT AACCTAACAG GCGGAGATGA CTGACGCCCT
32281 AGATCTAGAA ATGGACGGCA TCAGTACCGA GCAGCGTCTC CTAGAGAGGC GCAGGCAGGC
32341 GGCTGAGCAA GAGCGCCTCA ATCAGGAGCT CCGAGATCTC GTTAACCTGC ACCAGTGCAA
32401 AAGAGGCATC TTTTGTCTGG TAAAGCAGGC CAAAGTCACC TACGAGAAGA CCGGCAACAG
32461 CCACCGCCTC AGTTACAAAT TGCCCACCCA GCGCCAGAAG CTGGTGCTCA TGGTGGGTGA
32521 GAATCCCATC ACCGTCACCC AGCACTCGGT AGAGACCGAG GGGTGTCTGC ACTCTCCCTG
32581 TCGGGGTCCA GAAGACCTCT GCACCCTGGT AAAGACCCTG TGCGGTCTCA GAGATTTAGT
32641 CCCCTTTAAC TAATCAAACA CTGGAATCAA TAAAAAGAAT CACTTACTTA AAATCAGACA
32701 GCAGGTCTCT GTCCAGTTTA TTCAGCAGCA CCTCCTTCCC CTCCTCCCAA CTCTGGTACT
32761 CCAAACGCCT TCTGGCGGCA AACTTCCTCC ACACCCTGAA GGGAATGTCA GATTCTTGCT
32821 CCTGTCCCTC CGCACCCACT ATCTTCATGT TGTTGCAGAT GAAGCGCACC AAAACGTCTG
32881 ACGAGAGCTT CAACCCCGTG TACCCCTATG ACACGGAAAG CGGCCCTCCC TCCGTCCCTT
32941 TCCTCACCCC TCCCTTCGTG TCTCCCGATG GATTCCAAGA AAGCCCCCCC GGGGTCTGT
33001 CTCTGAACCT GGCCGAGCCC CTGGTCACTT CCCACGGCAT GCTCGCCCTG AAAATGGGAA

Fig. 5S

SEQ ID NO:1

24/153

33061 GTGGCCTCTC CCTGGACGAC GCTGGCAACC TCACCTCTCA AGATATCACC ACCGCTAGCC
33121 CTCCCCCTCA AAAAACCAAG ACCAACCTCA GCCTAGAAAC CTCATCCCCC CTAAGTGTAA
33181 GCACCTCAGG CGCCCTCACC GTAGCAGCCG CCGCTCCCCT GGCAGTGGCC GGCACCTCCC
33241 TCACCATGCA ATCAGAGGCC CCCCTGACAG TACAGGATGC AAAACTCACC CTGGCCACCA
33301 AAGGCCCCCT GACCGTGTCT GAAGGCAAAC TGGCCTTGCA AACATCGGCC CCGCTGACGG
33361 CCGCTGACAG CAGCACCCCTC ACCGTTAGCG CCACACCACC AATTAATGTA AGCAGTGGAA
33421 GTTTAGGCTT AGACATGGAA GACCCTATGT ATACTCACGA TGGAAAAC TGGAATAAGAA
33481 TTGGGGGTCC ACTAAGAGTA GTAGACAGCT TGCACACACT CACTGTAGTT ACCGGAAATG
33541 GACTAACTGT AGATAACAAT GCCCTCCAAA CTAGAGTTAC GGGCGCCCTA GGTTATGACA
33601 CATCAGGAAA TCTACAATTG AGAGCTGCAG GAGGTATGCG AATTGATGCA AATGGCCAAC
33661 TTATCCTTAA TGTGGCATA CATTGATG CTCAGAACAA TCTCAGCCTT AGACTTGGTC
33721 AGGGACCCCT GTATATAAAC ACAGACCACA ACCTGGATTT GAATTGCAAC AGAGGTCTAA
33781 CCACAAC TACCAACAAC ACAAAAAAC TTGAGACTAA AATTAGCTCA GGCTTAGACT
33841 ATGACACCAA TGGTGCTGTC ATTATTAAAC TTGGCACTGG TCTAAGCTTC GACAACACAG
33901 GCGCCCTAAC TGTGGGAAAC ACTGGTGATG ATAACTGAC TCTGTGGACG ACCCCAGACC
33961 CATCTCCAAA TTGCAGAATT CACTCAGACA AAGACTGCAA GTTTACTCTA GTCCTAACTA
34021 AGTGTGGAAG CCAAATCCTG GCCTCTGTG CCGCCCTAGC GGTATCAGGA AATCTGGCTT
34081 CGATAACAGG CACCGTTGCC AGCGTTACCA TCTTTCTCAG ATTTGATCAG AATGGAGTGC
34141 TTATGGAAAA CTCCTCGCTA GACAGGCAGT ACTGGAACCT CAGAAATGGC AACTCAACTA
34201 ACGCTGCCCC CTACACCAAT GCAGTTGGGT TCATGCCAAA CCTCGCAGCA TACCCCAAAA
34261 CGCAAAGCCA GACTGCTAAA AACAACATTG TAAGTCAGGT TTACTTGAAT GGAGACAAAT
34321 CCAAACCCAT GACCCCTACC ATCACCCTCA ATGGAAC TGAATCCAGT GAAACTAGCC
34381 AGGTGAGTCA CTACTCCATG TCATTTACAT GGGCTTGGGA AAGTGGGCAA TATGCCACTG
34441 AAACCTTTGC CACCAACTCC TTCACCTTTT CTTACATTGC TGAACAATAA AAAGCATGAC
34501 ACTGATGTTT ATTTCTGATT CTTATTTTAT TATTTTCAA CACAACAAA TCATTCAAGT
34561 CATCTTCCA TCTTAGCTTA ATAGACACAG TAGCTTAATA GACCCAGTAG TGCAAAGCCC
34621 CATCTAGCT TATAGATCAG ACAGTGATAA TTAACCACCA CCACCACCAT ACCTTTTGAT
34681 TCAGGAAATC ATGATCATCA CAGGATCCTA GTCTTCAGGC CGCCCCCTCC CTCCCAAGAC
34741 ACAGAATACA CAGTCCTCTC CCCCCGACTG GCTTTAAATA ACACCATCTG GTTGGTCACA

Fig. 5T

SEQ ID NO:1

25/153

34801 GACATGTTCT TAGGGGTTAT ATTCCACACG GTCTCCTGCC GCGCCAGGCG CTCGTCGGTG
34861 ATGTTGATAA ACTCTCCCGG CAGCTCGCTC AAGTTCACGT CGCTGTCCAG CGGCTGAACC
34921 TCCGGCTGAC GCGATAACTG TGCGACCGGC TGCTGGACAA ACGGAGGCCG CGCCTACAAG
34981 GGGGTAGAGT CATAATCCTC GGTCAGGATA GGGCGGTGAT GCAGCAGCAG CGAGCGAAAC
35041 ATCTGCTGCC GCCGCCGCTC CGTCCGGCAG GAAAACAACA AGCCGGTGGT CTCCTCCGCG
35101 ATAATCCGCA CCGCCCGCAG CATCAGCTTC CTCGTTCTCC GCGCGCAGCA CCTCACCCCTG
35161 ATCTCGCTCA AGTCGGCGCA GTAGGTACAG CACAGCACCA CGATGTTATT CATGATCCCA
35221 CAGTGCAGGG CGCTGTATCC AAAGCTCATG CCGGGAACCA CCGCCCCAC GTGGCCATCG
35281 TACCACAAGC GCACGTAAAT TAAGTGTGCA CCCCTCATGA ACGTGCTGGA CACAAACATT
35341 ACTTCCTTGG GCATGTTGTA ATTCACCACC TCCCGGTACC AGATAAACCT CTGGTTAAAC
35401 AGGGCACCTT CCACCACCAT CCTGAACCAA GAGGCCAGAA CCTGCCCACC GGCTATGCAC
35461 TGCAGGGAAC CCGGGTTGGA ACAATGACAA TGCAGACTCC AAGGCTCGTA ACCGTGGATC
35521 ATCCGGCTGC TGAAGGCATC GATGTTGGCA CAACACAGAC ACACGTGCAT GCACTTTCTC
35581 ATGATTAGCA GCTCTTCCCT CGTCAGGATC ATATCCCAAG GAATAACCCA TTCTTGAATC
35641 AACGTAAAC CCACACAGCA GGAAGGCCT CGCACATAAC TCACGTTGTG CATGGTCAGC
35701 GTGTTGCATT CTGGAACAG CGGATGATCC TCCAGTATCG AGGCGCGGGT CTCCTTCTCA
35761 CAGGGAGGTA AAGGGTCCCT GCTGTACGGA CTGCGCCGGG ACGACCGAGA TCGTGTGAG
35821 CGTAGTGTC TGGAAAAGG AACGCCGAC GTGGTCATAC TTCTTGAAGC AGAACCAGGT
35881 TCGCGCGTGG CAGGCCTCCT TGCGTCTGCG GTCTCGCCGT CTAGCTCGCT CCGTGTGATA
35941 GTTGTAGTAC AGCCACTCCC GCAGAGCGTC GAGGCGCACC CTGGCTTCCG GATCTATGTA
36001 GACTCCGTCT TGCACCGCGG CCCTGATAAT ATCCACCACC GTAGAATAAG CAACACCCAG
36061 CCAAGCAATA CACTCGCTCT GCGAGCGGCA GACAGGAGGA GCGGGCAGAG ATGGGAGAAC
36121 CATGATAAAA AACTTTTTTT AAAGAATATT TTCCAATTCT TCGAAAGTAA GATCTATCAA
36181 GTGGCAGCGC TCCCCTCCAC TGGCGCGGTC AAACCTTACG GCCAAAGCAC AGACAACGGC
36241 ATTTCTAAGA TGTTCCTTAA TGGCGTCCAA AAGACACACC GCTCTCAAGT TGCAGTAAAC
36301 TATGAATGAA AACCCATCCG GCTGATTTTC CAATATAGAC GCGCCGGCGG CGTCCACCAA
36361 ACCCAGATAA TTTTCTTCTC TCCAGCGGTT TAGAATCTGT CTAAGCAAAT CCCTTATATC
36421 AAGTCCGGCC ATGCCAAAAA TCTGCTCAAG AGCGCCCTCC ACCTTCATGA CCAAGCAGCG
36481 CATCATGATT GCAAAAATTC AGGTTCTTCA GAGACCTGTA TAAGATTCAA AATGGGAACA

Fig. 5U

ITR0048PV

SEQ ID NO:1

26/153

36541 TTAACAAAAA TTCCTCTGTC GCGCAGATCC CTTCGCAGGG CAAGCTGAAC ATAATCAGAC
36601 AGGTCTGAAC GGACCAGTGA GGCCAAATCC CCACCAGGAA CCAGATCCAG AGACCCTATA
36661 CTGATTATGA CGCGCATACT CGGGGCTATG CTGACCAGCG TAGCGCCGAT GTAGGCGTGC
36721 TGCATGGGCG GCGAGATAAA ATGCAAAGTG CTGGTTAAAA AATCAGGCAA AGCCTCGCGC
36781 AAAAAAGCTA ACACATCATA ATCATGCTCA TGCAGGTAGT TGCAGGTAAG CTCAGGAACC
36841 AAAACGGAAT AACACACGAT TTTCTCTCA AACATGACTT CGCGGATACT GCGTAAAAACA
36901 AAAATTATAA ATAAAAAATT AATTAACTTA AACATTGGAA GCCTGTCTCA CAACAGGAAA
36961 AACCACCTTA ATCAACATAA GACGGGCCAC GGGCATGCCG GCATAGCCGT AAAAAAATTG
37021 GTCCCCGTGA TTAACAAGTA CCACAGACAG CTCCCCGGTC ATGTCGGGGG TCATCATGTG
37081 AGACTCTGTA TACACGTC TGATTGTGAAC ATCAGACAAA CAAAGAAATC GAGCCACGTA
37141 GCCCGGAGGT ATAATCACCC GCAGGCGGAG GTACAGCAAA ACGACCCCCA TAGGAGGAAT
37201 CACAAAATTA GTAGGAGAAA AAAATACATA AACACCAGAA AAACCCTGTT GCTGAGGCAA
37261 AATAGCGCCC TCCCGATCCA AAACAACATA AAGCGCTTCC ACAGGAGCAG CCATAACAAA
37321 GACCCGAGTC TTACCAGTAA AAGAAAAAAG ATCTCTCAAC GCAGCACCAG CACCAACACT
37381 TCGCAGTGTA AAAGGCCAAG TGCCGAGAGA GTATATATAG GAATAAAAAG TGACGTAAAC
37441 GGGCAAAGTC CAAAAAACGC CCAGAAAAAC CGCACGCGAA CCTACGCCCC GAAACGAAAG
37501 CCAAAAAACA CTAGACACTC CCTTCCGGCG TCAACTTCCG CTTTCCACG CTACGTCACT
37561 TGCCCCAGTC AAACAAACTA CATATCCCGA ACTTCCAAGT CGCCACGCCC AAAACACCGC
37621 CTACACCTCC CCGCCCGCCG GCCCGCCCC AAACCCGCCT CCCGCCCCGC GCCCCGCCTC
37681 GCGCCGCCCA TCTCATTATC ATATTGGCTT CAATCCAAAA TAAGGTATAT TATTGATGAT
37741 G

Fig. 5V

SEQ ID NO:2

27/153

1 CATCATCAAT AATATACCTC AAACCTTTTGG TGC GCGTTAA TATGCAAATG AGCCGTTTGA
61 ATTTGGGGAT GCGGGGCGCT GATTGGCTGC GGGAGCGGCG ACCGTTAGGG GCGGGGCGGG
121 TGACGTTTTG ATGACGTGTT TGTGAGGCGG AGCCGTTTGA CAAGTTCTCG TGGGAAAAGT
181 GACGTCAAAC GAGGTGTGGT TTGAACACGG AAATACTCAA TTTTCCCGCG CTCTCTGACA
241 GGAAATGAGG TGTTTCTGGG CGGATGCAAG TGAAAACGGG CCATTTTTCGC GCGAAAAC TG
301 AATGAGGAAG TGAAAATCTG AGTAATTTTCG CGTTTATGGC AGGGAGGAGT ATTTGCCGAG
361 GGCCGAGTAG ACTTTGACCG ATTACGTGGG GGTTCGATT ACCGTATTTT TCACCTAAAT
421 TTCCGCGTAC GGTGTCAAAG TCCGGTGT TTACGTAGGC GTCAGCTGAT CGCCAGGGTA
481 TTTAAACCTG CGCTCTCTAG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGTTTCT
541 CCTCCGCGCC GCGAGTCAGA TCTACACTTT GAAAGATGAG GCACCTGAGA GACCTGCCCG
601 GTAATGTTTT CCTGGCTACT GGGAACGAGA TTCTGGAAC TGTGGTGGAC GCCATGATGG
661 GTGACGACCC TCCTGAGCCC CCTACCCCAT TTGAGGCGCC TTCGCTGTAC GATTGTATG
721 ATCTGGAGGT GGATGTGCCC GAGAACGACC CCAACGAGGA GGCGGTGAAT GATTGTTTA
781 GCGATGCCGC GCTGCTGGCC GCCGAGCAGG CTAATACGGA CTCTGGCTCA GACAGCGATT
841 CCTCTCTCCA TACCCCGAGA CCCGGCAGAG GTGAGAAAAA GATCCCGAG CTAAAGGGG
901 AAGAGCTCGA CCTGCGCTGC TATGAGGAAT GCTTGCCTCC GAGCGATGAT GAGGAGGACG
961 AGGAGGCGAT TCGAGCTGCA GCGAGCGAGG GAGTGAAAGC TGCGGGCGAG AGCTTTAGCC
1021 TGGACTGTCC TACTCTGCCC GGACACGGCT GTAAGTCTTG TGAATTTTAT CGCATGAATA
1081 CTGGAGATAA GAATGTGATG TGTGCCCTGT GCTATATGAG AGCTTACAAC CATTGTGTTT
1141 ACAGTAAGTG TGATTAACCT TAGCTGGGAA GGCAGAGGGT GACTGGGTGC TGACTGGTTT
1201 ATTTATGTAT ATGTTTTTTA TGTGTAGGTC CCGTCTCTGA CGTAGATGAG ACCCCACTT
1261 CAGAGTG CAT TTCATACCC CCAGAAATTG GCGAGGAACC GCCCGAAGAT ATTATTCATA
1321 GACCAGTTGC AGTGAGAGTC ACCGGGCGGA GAGCAGCTGT GGAGAGTTTG GATGACTTGC
1381 TACAGGGTGG GGATGAACCT TTGGACTTGT GTACCCGGA ACGCCCCAGG CACTAAGTGC
1441 CACACATGTG TGTTTACTTA AGGTGATGTC AGTATTTATA GGGTGTGGAG TGCAATAAAA
1501 TCCGTGTTGA CTTTAAGTGT GTGGTTTATG ACTCAGGGGT GGGGACTGTG GGTATATAAG
1561 CAGGTGCAGA CCTGTGTGGT CAGTTCAGAG CAGGACTCAT GGAGATCTGG ACGGTCTTGG
1621 AAGACTTTCA CCAGACTAGA CAGCTGCTAG AGAACTCATC GGAGGAAGTC TCTTACCTGT
1681 GGAGATTTTG CTTCGGTGGG GCTCTAGCTA AGCTAGTCTA TAGGGCCAAA CAGGATTATA

Fig. 6A

SEQ ID NO:2

28/153

1741 AGGATCAATT TGAGGATATT TTGAGAGAGT GTCCTAGTAT TTTTGACTCT CTCAACTTGG
1801 GCCATCAGTC TCACTTTAAC CAGAGTATTC TGAGAGCCCT TGACTTTTCT ACTCCTGGCA
1861 GAACTACCGC CGCGGTAGCC TTTTTTGCCT TTATTCTTGA CAAATGGAGT CAAGAAACCC
1921 ATTTTCAGCAG GGATTACCGT CTGGACTGCT TAGCAGTAGC TTTGTGGAGA ACATGGAGGT
1981 GCCAGCGCCT GAATGCAATC TCCGGCTACT TGCCAGTACA GCCGGTAGAC ACGCTGAGGA
2041 TCCTGAGTCT CCAGTCACCC CAGGAACACC AACGCCGCCA GCAGCCGCAG CAGGAGCAGC
2101 AGCAAGAGGA GGAGGAGGAG GAGGACCGAG AAGAGAACCC GAGAGCCGGT CTGGACCCTC
2161 CGGTGGCGGA GGAGGAGGAG TAGCTGACTT GTTTCCCGAG CTGCGCCGGG TGCTGACTAG
2221 GTCTTCCAGT GGACGGGAGA GGGGGATTAA GCGGGAGAGG CATGAGGAGA CTAGTCACAG
2281 AACTGAACTG ACTGTCAGTC TGATGAGCCG CAGGCGCCCA GAATCGGTGT GGTGGCATGA
2341 GGTTCAGTCG CAGGGGATAG ATGAGGTCTC GGTAATGCAT GAGAAATATT CCCTAGAACA
2401 AGTCAAGACT TGTGTTGG AGCCCGAGGA TGATTGGGAG GTAGCCATCA GGAATTATGC
2461 CAAGCTGGCT CTGAGGCCAG ACAAGAAGTA CAAGATTACC AAAGTGATTA ATATCAGAAA
2521 TTCCTGCTAC ATTTCCGGGA ATGGGGCCGA GGTGGAGATC AGTACCCAGG AGAGGGTGGC
2581 CTTTCAGATGT TGTATGATGA ATATGTACCC GGGGGTGGTG GGCATGGAGG GAGTCACCTT
2641 TATGAACCGC AGGTTTAGGG GTGATGGGTA TAATGGGGTG GTCTTTATGG CCAACACCAA
2701 GCTGACAGTG CACGGATGCT CCTTCTTTGG CTTCAATAAC ATGTGCATCG AGGCCTGGGG
2761 CAGTGTTCAT GTGAGGGGAT GCAGCTTTTC AGCCAAGTGG ATGGGGGTCTG TGGGCAGAAC
2821 CAAGAGCGTG GTGTCAGTGA AGAAATGCCT GTTCGAGAGG TGCCACCTGG GGGTGATGAG
2881 CGAGGGCGAA GCCAAAGTCA AACACTGCGC CTCTACCGAG ACGGGCTGCT TTGTGATGAT
2941 CAAGGGCAAT GCCAAAGTCA AGCATAACAT GATTTGTGGG GCCTCGGATG AGCGCGGCTA
3001 CCAGATGCTG ACCTGTGCCG GTGGGAACAG CCATATGCTG GCCACCGTGC ATGTGGCCTC
3061 GCACCCCCGC AAGACATGGC CCGAGTTCGA GCACAACGTC ATGACCCGCT GCAATGTGCA
3121 CCTGGGGTCC CGCCGAGGCA TGTTTCATGCC CTACCAGTGC AACATGCAAT TTGTGAAGGT
3181 GCTGCTGGAG CCCGATGCCA TGTCCAGAGT GAGCCTGGTG GGGGTGTTTG ACATGAATGT
3241 GGAGGTGTGG AAAATTCTGA GATATGATGA ATCCAAGACC AGGTGCCGGG CCTGCGAATG
3301 CGGAGGCAAG CACGCCAGGC TTCAGCCCGT GTGTGTGGAG GTGACGGAGG ACCTGCGACC
3361 CGATCATTTG GTGTTGTCCT GCAACGGGAC GGAGTTCGGC TCCAGCGGGG AAGAATCTGA
3421 CTAGAGTGAG TAGTGTGTTG GGGTGGGTGG GAGTCTGCAT GATGGGCAGA ATGACTAAAA

Fig. 6B

SEQ ID NO:2

29/153

3481 TCTGTGTTTT TCTGCGCAGC AGCATGAGCG GAAGCGCCTC CTTTGAGGGA GGGGTATTCA
3541 GCCCTTATCT GACGGGGCGT CTCCCCTCCT GGGCGGGAGT GCGTCAGAAT GTGATGGGAT
3601 CCACGGTGGA CGGCCGGCCC GTGCAGCCCG CGAACTCTTC AACCTGACC TACGCGACCC
3661 TGAGCTCCTC GTCCGTGGAC GCAGCTGCCG CCGCAGCTGC TGCTTCCGCC GCCAGCGCCG
3721 TGCGCGGAAT GGCTTGGGC GCCGGCTACT ACAGCTCTCT GGTGGCCAAC TCGAGTTCCA
3781 CCAATAATCC CGCCAGCCTG AACGAGGAGA AGCTGCTGCT GCTGATGGCC CAGCTCGAGG
3841 CCCTGACCCA GCGCCTGGGC GAGCTGACCC AGCAGGTGGC TCAGCTGCAG GCGGAGACGC
3901 GGGCCGCGGT TGCCACGGTG AAAACCAAAT AAAAAATGAA TCAATAAATA AACGGAGACG
3961 GTTGTGATT TTAACACAGA GTCTTGATCT TTATTTGATT TTTCGCGCGC GGTAGGCCCT
4021 GGACCACCGG TCTCGATCAT TGAGCACCCG GTGGATTTTT TCCAGGACCC GGTAGAGGTG
4081 GGCTTGATG TTGAGGTACA TGGGCATGAG CCCGTCCCGG GGGTGGAGGT AGCTCCATTG
4141 CAGGGCCTCG TGCTCGGGG TGGTGTGTA AATCACCCAG TCATAGCAGG GGCGCAGGGC
4201 GTGGTGCTGC ACGATGTCCT TGAGGAGGAG ACTGATGGCC ACGGGCAGCC CCTTGGTGTA
4261 GGTGTGACG AACCTGTTGA GCTGGGAGGG ATGCATGCGG GGGGAGATGA GATGCATCTT
4321 GGCCTGGATC TTGAGATTGG CGATGTTCCC GCCCAGATCC CGCCGGGGGT TCATGTTGTG
4381 CAGGACCACC AGCACGGTGT ATCCGGTGCA CTTGGGGAAT TTGTCATGCA ACTTGAAGG
4441 GAAGGCGTGA AAGAATTTGG AGACGCCCTT GTGACCGCCC AGGTTTCCA TGCACTCATC
4501 CATGATGATG GCGATGGGCC CGTGGGCGGC GGCCTGGGCA AAGACGTTTC GGGGGTCGGA
4561 CACATCGTAG TTGTGGTCCT GGGTGAGCTC GTCATAGGCC ATTTTAATGA ATTTGGGGCG
4621 GAGAGTGCCC GACTGGGGGA CGAAGGTGCC CTCGATCCCG GGGGCGTAGT TCCCCTCGCA
4681 GATCTGCATC TCCCAGGCCT TGAGCTCGGA GGGGGGATC ATGTCCACCT GCGGGGCGAT
4741 GAAAAAACC GTTTCCGGGG CGGGGGAGAT GAGCTGGGCC GAAAGCAGGT TCCGGAGCAG
4801 CTGGGACTTG CCGCAGCCGG TGGGACCGTA GATGACCCCG ATGACCGGCT GCAGGTGGTA
4861 GTTGAGGGAG AGACAGCTGC CATCCTCGCG GAGGAGGGGG GCCACCTCGT TCATCATCTC
4921 GCGCACATGC ATGTTCTCGC GCACGAGTTC CGCCAGGAGG CGCTCGCCCC CCAGCGAGAG
4981 GAGCTCTTGC AGCGAGGCGA AGTTTTTCAG CGGCTTGAGC CCGTCGGCCA TGGGCATTTT
5041 GGAGAGGGTC TGTTGCAAGA GTTCAGACG GTCCCAGAGC TCGGTGATGT GCTCTAGGGC
5101 ATCTCGATCC AGCAGACCTC CTCGTTTCGC GGGTTGGGGC GACTGCGGGA GTAGGGCACC
5161 AGGCGATGGG CGTCCAGCGA GGCCAGGGTC CGGTCCTTCC AGGGTCGCAG GGTCCGCGTC

Fig. 6C

SEQ ID NO:2

30/153

5221 AGCGTGGTCT CCGTCACGGT GAAGGGGTGC GCGCCGGGCT GGGCGCTTGC GAGGGTGCGC
5281 TTCAGGCTCA TCCGGCTGGT CGAGAACCGC TCCC GGTCGG CGCCCTGCGC GTCGGCCAGG
5341 TAGCAATTGA GCATGAGTTC GTAGTTGAGC GCCTCGGCCG CGTGGCCCTT GGC CGGAGC
5401 TTACCTTTGG AAGTGTGTCC GCAGACGGGA CAGAGGAGGG ACTTGAGGGC GTAGAGCTTG
5461 GGGGCGAGGA AGACGGACTC GGGGGCGTAG GCGTCCGCGC CGCAGCTGGC GCAGACGGTC
5521 TCGCACTCCA CGAGCCAGGT GAGGTCGGGG CGGTCGGGGT CAAAAACGAG GTTTCCTCCG
5581 TGCTTTTTGA TGC GTTTCTT ACCTCTGGTC TCCATGAGCT CGTG TCCCCG CTGGGTGACA
5641 AAGAGGCTGT CCGTGTCCCC GTAGACCGAC TTTATGGGCC GGTCTCGAG CGGGGTGCCG
5701 CGGTCTCGT CGTAGAGGAA CCCC GCCAC TCCGAGACGA AGGCCCGGGT CCAGGCCAGC
5761 ACGAAGGAGG CCACGTGGA GGGGTAGCGG TC GTTGCCA CCAGCGGGTC CACCTTCTCC
5821 AGGGTATGCA AGCACATGTC CCCCTCGTCC ACATCCAGGA AGGTGATTGG CTTGTAAGTG
5881 TAGGCCACGT GACCGGGGGT CCGGCCGGG GGGGTATAAA AGGGGGCGGG CCCCTGCTCG
5941 TCCTCACTGT CTCCGGATC GCTGTCCAGG AGCGCCAGCT GTTGGGGTAG GTATTCCCTC
6001 TCGAAGGCGG GCATGACCTC GGC ACTCAGG TTGT CAGTTT CTAGAAACGA GGAGGATTTG
6061 ATATTGACGG TGCCGTGGA GACGCCTTTC ATGAGCCCCT CGTCCATCTG GTCAGAAAAG
6121 ACGATCTTTT TGTGTCGAG CTTGGTGCGG AAGGAGCCGT AGAGGGCGTT GGAGAGGAGC
6181 TTGGCGATGG AGCGCATGGT CTG GTTCTT TCCTTGTCGG CGCGCTCCTT GGC GGCGATG
6241 TTGAGCTGCA CGTACTCGCG CGCCACGCAC TTCCATTCTG GGAAGACGGT GGTGAGCTCG
6301 TCGGGCACGA TTCTGACCCG CCAGCCGCGG TTGTGCAGGG TGATGAGGTC CACGCTGGTG
6361 GCCACCTCGC CGCGCAGGG CTC GTTGGTC CAGCAGAGGC GCGGCCCTT GCGCGAGCAG
6421 AAGGGGGGCA GCGGGTCCAG CATGAGCTCG TCTGGGGGGT CGGCGTCCAC GGTGAAGATG
6481 CCGGGCAGGA GCTCGGGGTC GAAGTAGCTG ATGGAAGTGG CCAGATCGTC CAGGGAAGCT
6541 TGCCAGTCGC GCACGGCCAG CGCGCGCTCG TAGGGGCTGA GGGGCGTGCC CCAGGGCATG
6601 GGGTGCGTGA GCGCGGAGGC GTACATGCCG CAGATGTCGT AGACGTAGAG GGGCTCCTCG
6661 AGGATGCCGA TG TAGGTGGG GTAGCAGCG CCCCCGCGGA TGCTGGCGCG CACGTAGTCG
6721 TACAGCTCGT GCGAGGGCGC GAGGAGCCCC GTGCCGAGAT TGGAGCGCTG CGGCTTTTCG
6781 GCGCGGTAGA CGATCTGGCG GAAGATGGCG TGGGAGTTGG AGGAGATGGT GGCCTCTGG
6841 AAGATGTTGA AGTGGGCGTG GGCAGGCCG ACCGAGTCCC TGATGAAGTG GCGTAGGAG
6901 TCCTGCAGCT TGGCGACGAG CTCGGCGGTG ACGAGGACGT CCAGGGCGCA GTAGTCGAGG

Fig. 6D

SEQ ID NO:2

31/153

6961 GTCTCTTGA TGATGTCATA CTTGAGCTGG CCCTTCTGCT TCCACAGCTC GCGGTTGAGA
7021 AGGAACTCTT CGCGGTCCTT CCAGTACTCT TCGAGGGGGA ACCCGTCCTG ATCGGCACGG
7081 TAAGAGCCCA CCATGTAGAA CTGGTTGACG GCCTTGTAAG CGCAGCAGCC CTTCTCCACG
7141 GGGAGGGCGT AAGCTTGCGC GGCCTTGCGC AGGGAGGTGT GGGTGAGGGC GAAGGTGTCTG
7201 CGCACCATGA CCTTGAGGAA CTGGTGCTTG AAGTCGAGGT CGTCGCAGCC GCCCTGCTCC
7261 CAGAGTTGGA AGTCCGTGCG CTTCTTGTAAG GCGGGGTTGG GCAAAGCGAA AGTAACATCG
7321 TTGAAGAGGA TCTTGCCCGC GCGGGGCATG AAGTTGCGAG TGATGCGGAA AGGCTGGGGC
7381 ACCTCGGCCC GGTGTGTGAT GACCTGGGCG GCGAGGACGA TCTCGTCGAA GCCGTTGATG
7441 TTGTGCCCCG CGATGTAGAG TTCCACGAAT CGCGGGCGGC CCTTGACGTG GGGCAGCTTC
7501 TTGAGCTCGT CGTAGGTGAG CTCGGCGGGG TCGCTGAGCC CGTGCTGTTC GAGGGCCAG
7561 TCGGCGACGT GGGGTTGGC GCTGAGGAAG GAAGTCCAGA GATCCACGGC CAGGGCGGTC
7621 TGCAAGCGGT CCCGGTACTG ACGGAAGTGC TGGCCACCGG CCATTTTTTC GGGGGTGACG
7681 CAGTAGAAGG TGCGGGGGTC GCCGTGCCAG CGGTCCCACT TGAGCTGGAG GGCGAGGTCTG
7741 TGGGCGAGCT CGACGAGCGG TGGGTCCCCG GAGAGTTTCA TGACCAGCAT GAAGGGGACG
7801 AGCTGCTTGC CGAAGGACCC CATCCAGGTG TAGGTTTCCA CATCGTAGGT GAGGAAGAGC
7861 CTTTCGGTGC GAGGATGCGA GCCGATGGGG AAGAACTGGA TCTCCTGCCA CCAGTTGGAG
7921 GAATGGCTGT TGATGTGATG GAAGTAGAAA TGCCGACGGC GCGCCGAGCA CTCGTGCTTG
7981 TGTTTATACA AGCGTCCGCA GTGCTCGCAA CGCTGCACGG GATGCACGTG CTGCACGAGC
8041 TGTACCTGAG TTCCTTTGAC GAGGAATTTT AGTGGGCAGT GGAGCGCTGG CGGCTGCATC
8101 TGGTGCTGTA CTACGTCTTG GCCATCGGCG TGGCCATCGT CTGCCCTCGAT GGTGGTCATG
8161 CTGACGAGGC CGCGCGGGAG GCAGGTCCAG ACCTCGGCTC GGACGGGTCTG GAGAGCGAGG
8221 ACGAGGGCGC GCAGGCCGGA GCTGTCCAGG GTCCTGAGAC GCTGCGGAGT CAGGTCAGTG
8281 GGCAGCGGCG GCGCGCGGTT GACTTGCAAG AGCTTTTCCA GGGCGCGCGG GAGGTCCAGA
8341 TGGTACTTGA TCTCCACGGC GCCGTTGGTG GCGACGTCCA CGGCTTGCAAG GGTCCCGTGC
8401 CCCTGGGGCG CCACCACCGT GCGGCGTTTC TTCTTGGGCG CTGGTTCCAT GCCGGTCAGA
8461 AGCGGCGGCG AGGACGCGCG CCGGGCGGCA GGGGCGGCTC GGGGCCCCGA GGCAGGGGCG
8521 GCAGGGGCAC GTCGGCGCCG CGCGCGGGCA GGTCTGGTA CTGCGCCCGG AGAAGACTGG
8581 CGTGAGCGAC GACGCGACGG TTGACGTCTT GGATCTGACG CCTCTGGGTG AAGGCCACGG
8641 GACCCGTGAG TTTGAACCTG AAAGAGAGTT CGACAGAATC AATCTCGGTA TCGTTGACGG

Fig. 6E

SEQ ID NO:2

32/153

8701 CGGCCTGCCG CAGGATCTCT TGCACGTCGC CCGAGTTGTC CTGGTAGGCG ATCTCGGTCA
8761 TGAAGTCTCT GATCTCTCTC TCCTGAAGGT CTCCGCGGCC GGCGCGCTCG ACGGTGGCCG
8821 CGAGGTCGTT GGAGATGCGG CCCATGAGCT GCGAGAAGGC GTTCATGCCG GCCTCGTTCC
8881 AGACGCGGCT GTAGACCACG GCTCCGTCGG GGTGCGCGCG GCGCATGACC ACCTGGGCGA
8941 GGTGAGCTC GACGTGGCGC GTGAAGACCG CGTAGTTGCA GAGGCGCTGG TAGAGGTAGT
9001 TGAGCGTGGT GGCGATGTGC TCGGTGACGA AGAAGTACAT GATCCAGCGG CGGAGCGGCA
9061 TCTCGCTGAC GTCGCCCAGG GCTTCCAAGC GCTCCATGGC CTCGTAGAAG TCCACGGCGA
9121 AGTTGAAAAA CTGGGAGTTG CGCGCCGAGA CCGTCAACTC CTCTCCAGA AGACGGATGA
9181 GCTCGGCGAT GGTGGCGCGC ACCTCGCGCT CGAAGGCCCC GGGGGGCTCC TCTTCTTCCA
9241 TCTCCTCCTC CTCTTCTCTC TCCACTAACA TCTCTCTAC TTCTCCTCA GGAGGCGGTG
9301 GCGGGGGAGG GGCCCTGCGT CGCCGGCGGC GCACGGGCAG ACGGTCGATG AAGCGCTCGA
9361 TGGTCTCCCC GCGCCGGCGA CGCATGGTCT CCGTGACGGC GCGCCCGTCC TCGCGGGGCC
9421 GCAGCGTGAA GACGCCGCCG CGCATCTCCA GGTGGCCGCC GGGGGGGTCT CCGTTGGGCA
9481 GGGAGAGGGC GCTGACGATG CATCTTATCA ATTGGCCCGT AGGGACTCCG CGCAAGGACC
9541 TGAGCGTCTC GAGATCCACG GGATCCGAAA ACCGCTGAAC GAAGGCTTCG AGCCAGTCGC
9601 AGTCGCAAGG TAGGCTGAGC CCGGTTTCTT CGGGTATTG GTCGGGAGGC GGGCGGGCGA
9661 TGCTGCTGGT GATGAAGTTG AAGTAGGCGG TCCTGAGACG GCGGATGGTG GCGAGGAGCA
9721 CCAGGTCCTT GGGCCCGGCT TGCTGGATGC GCAGACGGTC GGCCATGCCC CAGGCGTGGT
9781 CCTGACACCT GGCGAGGTCC TTGTAGTAGT CCTGCATGAG CCGCTCCACG GGCACCTCCT
9841 CCTCGCCCGC GCGGCCGTGC ATGCGCGTGA GCCCGAACC GCGCTGCGGC TGGACGAGCG
9901 CCAGGTCGGC GACGACGCGC TCGGCGAGGA TGGCCTGCTG GATCTGGGTG AGGGTGGTCT
9961 GGAAGTCGTC GAAGTCGACG AAGCGGTGGT AGGCTCCGGT GTTGATGGTG TATGAGCAGT
10021 TGGCCATGAC GGACCAGTTG ACGGTCCTGGT GGCCGGGGCG CACGAGCTCG TGGTACTTGA
10081 GCGCGGAGTA GCGCGCGGTG TCGAAGATGT ACTCGTTGCA GGTGCGCACG AGGTACTGGT
10141 ATCCGACGAG GAAGTGCGGC GCGGGCTGGC GGTAGAGCGG CCATCGCTCG GTGGCGGGGG
10201 CGCCGGGCGC GAGGTCCTCG AGCATGAGGC GGTGGTAGCC GTAGATGTAC CTGGACATCC
10261 AGGTGATGCC GCGGCGGGTG GTGGAGGCGC GCGGGAATC GCGGACGCGG TTCCAGATGT
10321 TGCGCAGCGG CAGGAAGTAG TTCATGGTGG CCGCGGTCTG GCCCGTGAGG CGCGCGCAGT
10381 CGTGGATGCT CTATACGGGC AAAAACGAAA GCGGTCAGCG GCTCGACTCC GTGGCCTGGA

Fig. 6F

SEQ ID NO:2

33/153

10441 GGCTAAGCGA ACGGGTTGGG CTGCGCGTGT ACCCCGGTTC GAATCTCGAA TCAGGCTGGA
10501 GCCGCAGCTA ACGTGGTACT GGCAC TCCCG TCTCGACCCA AGCCTGCACA AAACCTCCAG
10561 GATACGGAGG CGGGTCGTTT TGCAACTTTT TGAGGCCGGA AATGAACTA GTAAGCGCGA
10621 AAAGCGGCCG ACCGCGATGG CTCGCTGCCG TAGTCTGGAG AAGAATCGCC AGGGTTGCGT
10681 TGCGGTGTGC CCCGGTTCGA GGCCGGCCGG ATTCCGCGGC TAACGAGGGC GTGGCTGCCC
10741 CGTCGTTTCC AAGACCCCTA GCCAGCCGAC TTCTCCAGTT ACGGAGCGAG CCCCTCTTTT
10801 GTTTTGTTTG TTTTGGCCAG ATGCATCCCG TACTGCGGCA GATGCGCCCC CACCACCCTC
10861 CACCGCAACA ACAGCCCACT CCACAGCCGG CGTTCTGCC CCCGCCCCAG CAGCAGCAAC
10921 TTCCAGCCAC GACCGCCGCG GCCGCCGTGA GCGGGGCTGG ACAGACTTCT CAGTATGACC
10981 ACCTGGCCTT GGAAGAGGGC GAGGGGCTGG CGCGCCTGGG GGCCTCGTCG CCGGAGCGGC
11041 ACCCGCGCGT GCAGATGAAA CGGGACGCTC GCGAGGCCTA CGTGCCCAAG CAGAACCTGT
11101 TCAGAGACAG GAGCGGCGAG GAGCCCGAGG AGATGCGCGC GGCCCGGTTT CACGCGGGGC
11161 GGGAGCTGCG GCGCGGCCTG GACCGAAAGA GGGTGCTGAG GGACGAGGAT TTCGAGGCGG
11221 ACGAGCTGAC GGGGATCAGC CCCGCGCGCG CGCACGTGGC CGCGGCCAAC CTGGTCACGG
11281 CGTACGAGCA GACCGTGAAG GAGGAGAGCA ACTTCCAAA ATCCTTCAAC AACCACGTGC
11341 GCACCCTGAT CGCGCGCGAG GAGGTGACCC TGGGCCTGAT GCACCTGTGG GACCTGCTGG
11401 AGGCCATCGT GCAGAACCCC ACCAGCAAGC CGCTGACGGC GCAGCTGTTT CTGGTGGTGC
11461 AGCACAGTCG GGACAACGAG GCGTTCAGGG AGGCGCTGCT GAATATCACC GAGCCCGAGG
11521 GCCGCTGGCT CCTGGACCTG GTGAACATTC TGCAGAGCAT CGTGGTGCAG GAGCGCGGGC
11581 TGCCGCTGTC CGAGAAGCTG GCGGCCATCA ACTTCTCGGT GCTGAGTCTG GGCAAGTACT
11641 ACGCTAGGAA GATCTACAAG ACCCCGTACG TGCCCATAGA CAAGGAGGTG AAGATCGACG
11701 GGTTTTACAT GCGCATGACC CTGAAAGTGC TGACCCTGAG CGACGATCTG GGGGTGTACC
11761 GCAACGACAG GATGCACCGC GCGGTGAGCG CCAGCCGCCG GCGCGAGCTG AGCGACCAGG
11821 AGCTGATGCA CAGCCTGCAG CGGGCCCTGA CCGGGGCCGG GACCGAGGGG GAGAGCTACT
11881 TTGACATGGG CGCGGACCTG CGCTGGCAGC CCAGCCGCCG GGCCTTGGA GCTGCCGGCG
11941 GTTCCCCCTA CGTGGAGGAG GTGGACGATG AGGAGGAGGA GGGCGAGTAC CTGGAAGACT
12001 GATGGCGCGA CCGTATTTTT GCTAGATGCA GCAACAGCCA CCGCCGCCTC CTGATCCCGC
12061 GATGCGGGCG GCGCTGCAGA GCCAGCCGTC CGGCATTAAC TCCTCGGACG ATTGGACCCA
12121 GGCCATGCAA CGCATCATGG CGCTGACGAC CCGCAATCCC GAAGCCTTTA GACAGCAGCC

Fig. 6G

ITR0048PV

SEQ ID NO:2

34/153

12181 TCAGGCCAAC CGGCTCTCGG CCATCCTGGA GGCCGTGGTG CCCTCGCGCT CGAACCCAC
12241 GCACGAGAAG GTGCTGGCCA TCGTGAACGC GCTGGTGGAG AACAGGCCA TCCGCGGCGA
12301 CGAGGCCGGG CTGGTGTACA ACGCGCTGCT GGAGCGCGTG GCCCGCTACA ACAGCACCAA
12361 CGTGCAGACG AACCTGGACC GCATGGTGAC CGACGTGCGC GAGGCGGTGT CGCAGCGCGA
12421 GCGGTTCCAC CGCGAGTCGA ACCTGGGCTC CATGGTGGCG CTGAACGCCT TCCTGAGCAC
12481 GCAGCCCGCC AACGTGCCCC GGGGCCAGGA GGA CTACACC AACTTCATCA GCGCGCTGCG
12541 GCTGATGGTG GCCGAGGTGC CCCAGAGCGA GGTGTACCAG TCGGGGCCGG ACTACTTCTT
12601 CCAGACCAGT CGCCAGGGCT TGCAGACCGT GAACCTGAGC CAGGCTTTCA AGA ACTTGCA
12661 GGGACTGTGG GGC GTGCAGG CCCC GGTCGG GGACCGCGC ACGGTGTCGA GCCTGCTGAC
12721 GCCGAACTCG CGCCTGCTGC TGCTGCTGGT GCGGCCCTTC ACGGACAGCG GCAGCGTGAG
12781 CCGCGACTCG TACCTGGGCT ACCTGCTTAA CCTGTACCGC GAGGCCATCG GGCAGGCGCA
12841 CGTGGACGAG CAGACCTACC AGGAGATCAC CCACGTGAGC CGCGCGCTGG GCCAGGAGGA
12901 CCCGGGCAAC CTGGAGGCCA CCCTGAACTT CTGCTGACC AACC GGTCGC AGAAGATCCC
12961 GCCCCAGTAC GCGCTGAGCA CCGAGGAGGA GCGCATCTG CGCTACGTGC AGCAGAGCGT
13021 GGGGCTGTTC CTGATGCAGG AGGGGGCCAC GCCCAGCGCC GCGCTCGACA TGACCGCGCG
13081 CAACATGGAG CCCAGCATGT ACGCCCGCAA CCGCCCGTTC ATCAATAAGC TGATGGACTA
13141 CTTGCATCGG GCGGCCGCCA TGA ACTCGGA CTACTTTACC AACGCCATCT TGAACCCGCA
13201 CTGGCTCCCG CCGCCCGGGT TCTACACGGG CGAGTATGAC ATGCCCGACC CCAACGACGG
13261 GTTCTGTGG GATGACGTGG ACAGCAGCGT GTTCTCGCCG CGCCCCGCCA CCACCGTGTG
13321 GAAGAAAGAG GCGGGGGACC GCGGCCGCTC CTCGGCGCTG TCCGGTCGCG CGGGTGCTGC
13381 CGCGGCGGTG CCCGAGGCCG CCAGCCCCTT CCCGAGCCTG CCCTTTTCGC TGAACAGCGT
13441 GCGCAGCAGC GAGCTGGGAC GGCTGACGCG GCCGCGCCTG CTGGGCGAGG AGGAGTACCT
13501 GAACGACTCC TTGTTGAGGC CCGAGCGCGA GAAGAACTTC CCAATAACG GGATAGAGAG
13561 CCTGGTGGAC AAGATGAGCC GCTGGAAGAC GTACGCGCAC GAGCACAGGG ACGAGCCGCG
13621 AGCTAGCAGC AGCACCGGCG CCCGTAGACG CCAGCGGCAC GACAGGCAGC GGGGACTGGT
13681 GTGGGACGAT GAGGATTCCG CCGACGACAG CAGCGTGTG GACTTGGGTG GGAGTGGTGG
13741 TGGTAAACCG TTCGCTCACC TGCGCCCCCG TATCGGGCGC CTGATGTAAG AATCTGAAAA
13801 AATAAAAAAC GGTACTCACC AAGGCCATGG CGACCAGCGT GCGTTCTTCT CTGTTGTTTG
13861 TAGTAGTATG ATGAGGCGCG TGTACCCGGA GGGTCCTCCT CCCTCGTACG AGAGCGTGAT

Fig. 6H

SEQ ID NO:2

35/153

13921 GCAGCAGGCG GTGGCGGCGG CGATGCAGCC CCCGCTGGAG GCGCCTTACG TGCCCCCGCG
13981 GTACCTGGCG CCTACGGAGG GGCGGAACAG CATTCGTTAC TCGGAGCTGG CACCCTTGTA
14041 CGATACCACC CGGTTGTACC TGGTGGACAA CAAGTCGGCG GACATCGCCT CGTGAACTA
14101 CCAGAACGAC CACAGCAACT TCCTGACCAC CGTGGTGCAG AACAAAGATT TCACCCCCAC
14161 GGAGGCCAGC ACCCAGACCA TCAACTTTGA CGAGCGCTCG CGGTGGGGCG GCCAGCTGAA
14221 AACCATCATG CACACCAACA TGCCCAACGT GAACGAGTTC ATGTACAGCA ACAAGTTCAA
14281 GGCGCGGGTC ATGGTCTCGC GCAAGACCCC CAACGGGGTC GCGGTAGGGG ATGATTATGA
14341 TGGTAGTCAG GACGAGCTGA CCTACGAGTG GGTGGAGTTT GAGCTGCCCC AGGGCAACTT
14401 CTCGGTGACC ATGACCATCG ATCTGATGAA CAACGCCATC ATCGACAATT ACTTGCGGT
14461 GGGACGGCAG AACGGGGTGC TGGAGAGCGA CATCGGCGTG AAGTTCGACA CGCGCAACTT
14521 CCGGCTGGGC TGGGACCCCG TGACCGAGCT GGTGATGCCG GGCGTGACA CCAACGAGGC
14581 CTTCCACCCC GACATCGTCC TGCTGCCCCG CTGCGGCGTG GACTTCACCG AGAGCCGCCT
14641 CAGCAACCTG CTGGGCATCC GCAAGCGGCA GCCCTTCCAG GAGGGCTTCC AGATCCTGTA
14701 CGAGGACCTG GAGGGGGGCA ACATCCCCGC GCTCTTGGAT GTCGAAGCCT ATGAAGAAAG
14761 TAAGGAAAAA GCAGAGGCTG AGGCAACTGC AGCCGTGGCT ACTGCCGCTG TCACCGATGC
14821 AGATGCAGCT ACTACCAGGG GCGATACATT CGCCACTGTG GCTGAAGAAG CAGCCGCCGT
14881 AGCGGCGACC GATGATAGTG AAAGTAAGAT AGTCATCAAG CCGGTGGAGA AGGACAGCAA
14941 GAACAGGAGC TACAACGTTT TATCGGATGG AAAGAACACC GCCTACCGCA GCTGGTACCT
15001 GGCCTACAAC TACGGCGACC CCGAGAAGGG CGTGCGCTCC TGGACGCTGC TCACCACCTC
15061 GGACGTCACC TGCGGCGTGG AGCAAGTCTA CTGGTCGCTG CCCGACATGA TGCAAGACCC
15121 GGTCACCTTC CGCTCCACGC GTCAAGTTAG CAACTACCCG GTGGTGGGCG CCGAGCTCCT
15181 GCGCGTCTAC TCCAAGAGCT TCTTCAACGA GCAGGCCGTC TACTCGCAGC AGCTGCGCGC
15241 CTTACCTCG CTCACGCACG TCTTCAACCG CTTCCCCGAG AACCAGATCC TCGTCCGCCC
15301 GCGCGCGCCC ACCATTACCA CCGTCAGTGA AAACGTTCC TCTCTCACAG ATCACGGGAC
15361 CCTGCCGCTG CGCAGCAGTA TCCGGGGAGT CCAGCGCGTG ACCGTCACTG ACGCCAGACG
15421 CCGCACCTGC CCCTACGTCT ACAAGGCCCT GGGCGTAGTC GCGCCGCGCG TCCTCTCGAG
15481 CCGCACCTTC TAAAAATGT CCATTCTCAT CTCGCCCAGT AATAACACCG GTTGGGGCCT
15541 GCGCGCGCCC AGCAAGATGT ACGGAGGCGC TCGCCAACGC TCCACGCAAC ACCCGTGC
15601 CGTGC GCGG CACTTCCGCG CTCCCTGGGG CGCCCTCAAG GGCGCGTGC GCTCGCGCAC

Fig. 6I

SEQ ID NO:2

36/153

15661 CACCGTCGAC GACGTGATCG ACCAGGTGGT GGCCGACGCG CGCAACTACA CGCCCCGCCG
15721 CGCGCCCGTC TCCACCGTGG ACGCCGTCAT CGACAGCGTG GTGGCCGACG CGCGCCGGTA
15781 CGCCCCGCGC AAGAGCCGGC GCGGCGCAT CGCCCCGGCG CACCGGAGCA CCCCCCCAT
15841 GCGCGCGGCG CGAGCCTTGC TGCAGAGGCG CAGGCGCACG GGACGCAGGG CCATGCTCAG
15901 GGCGGCCAGA CGCGCGGCCCT CTGGCAGCAG CAGCGCCGGC AGGACCCGCA GACGCGCGGC
15961 CACGGCGGCG GCGGCGGCCA TCGCCAGCAT GTCCCGCCCG CGGCGCGGCA ACGTGTACTG
16021 GGTGCGCGAC GCCGCCACCG GTGTGCGCGT GCCCGTGCAG ACCCGCCCCC CTCGCACTTG
16081 AAGATGCTGA CTTCGCGATG TTGATGTGTC CCAGCGGCGA GGAGGATGTC CAAGCGCAAA
16141 TTCAAGGAAG AGATGCTCCA GGTTCATCGCG CCTGAGATCT ACGGCCCCGC GCGGCGGGTG
16201 AAGGAGGAAA GAAAGCCCCG CAACTGAAG CGGGTCAAAA AGGACAAAAA GGAGGAGGAA
16261 GATGTGGACG GACTGGTGA GTTTGTGCGC GAGTTCGCCC CCCGGCGGCG CGTGCAGTGG
16321 CGCGGGCGGA AAGTGAAACC GGTGCTGCGA CCCGGCACCA CCGTGGTCTT CACGCCCCGC
16381 GAGCGTTCCG GCTCCGCCCT CAAGCGCTCC TACGACGAGG TGTACGGGGA CGAGGACATC
16441 CTCGAGCAGG CGGCCGAGCG TCTGGGCGAG TTTGCTTACG GCAAGCGCAG CCGCCCCGCG
16501 CCCTTGAAAG AGGAGGCGGT GTCCATCCCG CTGGACCACG GCAACCCAC GCGGAGTCTG
16561 AAGCCGGTGA CCCTGCAGCA GGTGCTGCCG AGCGCGGCGC CGCGCCGGGG CTTCAAGCGC
16621 GAGGGCGGCG AGGATCTGTA CCCGACCATG CAGCTGATGG TGCCCAAGCG CCAGAAGCTG
16681 GAGGACGTGC TGGAGCACAT GAAGGTGGAC CCCGAGGTGC AGCCCGAGGT CAAGGTGCGG
16741 CCCATCAAGC AGGTGGCCCC GGGCTGGGC GTGCAGACCG TGGACATCAA GATCCCCACG
16801 GAGCCCATGG AAACGCAGAC CGAGCCCGTG AAGCCCAGCA CCAGCACCAT GGAGGTGCAG
16861 ACGGATCCCT GGATGCCGGC GCCGGCTTCC ACCACCACCA CTCGCCGAAG ACGCAAGTAC
16921 GGCGCGGCCA GCCTGCTGAT GCCCAACTAC GCGCTGCATC CTTCCATCAT CCCCACGCCG
16981 GGCTACCGCG GCACGCGCTT CTACCGCGGC TACAGCAGCC GCCGCAAGAC CACCACCCGC
17041 CGCCGCCGTC GTCGCACCCG CCGCAGCAGC ACCCGCACTT CCGCCGCCTT GGTGCGGAGA
17101 GTGTACCGCA GCGGGCGCGA GCCTCTGACC CTGCCGCGCG CGCGCTACCA CCCGAGCATC
17161 GCCATTTAAC TCTGCCGTCG CCTCCTACTT GCAGATATGG CCCTCACATG CCGCCTCCGC
17221 GTCCCCATTA CGGGCTACCG AGGAAGAAAAG CCGCGCCGTA GAAGGCTGAC GGGGAACGGG
17281 CTGCGTCGCC ATCACCACCG GCGGCGGCGC GCCATCAGCA AGCGGTTGGG GGGAGGCTTC
17341 CTGCCCCGCG TGATCCCCAT CATCGCCGCG GCGATCGGGG CGATCCCCCG CATAGCTTCC

Fig. 6J

SEQ ID NO:2

37/153

17401 GTGGCGGTGC AGGCCTCTCA GCGCCACTGA GACACAGCTT GGAAAATTTG TAATAAAAAA
17461 TGGACTGACG CTCCTGGTCC TGTGATGTGT GTTTTATAGAT GGAAGACATC AATTTTTCGT
17521 CCCTGGCACC GCGACACGGC ACGCGGCCGT TTATGGGCAC CTGGAGCGAC ATCGGCAACA
17581 GCCAACTGAA CGGGGGCGCC TTCAATTGGA GCAGTCTCTG GAGCGGGCTT AAGAATTTTCG
17641 GGTCCACGCT CAAAACCTAT GGCAACAAGG CGTGGAACAG CAGCACAGGG CAGGCGCTGA
17701 GGGAAAAGCT GAAAGAGCAG AACTTCCAGC AGAAGGTGGT CGATGGCCTG GCCTCGGGCA
17761 TCAACGGGGT GGTGGACCTG GCCAACCAGG CCGTGCAGAA ACAGATCAAC AGCCGCTTGG
17821 ACGCGGTCCC GCCCGCGGGG TCCGTGGAGA TGCCCCAGGT GGAGGAGGAG CTGCTCCCC
17881 TGGACAAGCG CGGCGACAAG CGACCGCGTC CCGATGCAGA GGAGACGCTG CTGACGCACA
17941 CGGACGAGCC GCCCCGTAC GAGGAGGCGG TGAAACTGGG TCTGCCCACC ACGCGGCCCG
18001 TGGCGCTCTT GGCCACCGGG GTGCTGAAAC CCAGCAGCAG CAGCCAGCCC GCGACCCTGG
18061 ACTTGCTTCC GCCTGCTTCC CGCCCCCTCA CAGTGGCTAA GCCCTGCCG CCGGTGGCCG
18121 TCGCGTCGCG CGCCCCCGA GGCGCCCCC AGGCGAACTG GCAGAGCACT CTGAACAGCA
18181 TCGTGGGTCT GGGAGTGCAG AGTGTGAAGC GCCGCCGCTG CTATTAAAAG AACTGTAGC
18241 GCTTAACTTG CTTGTCTGTG TGTATATGTA TGTCCGCCGA CCAGAAGGAA GAGGCGCGTC
18301 GCCGAGTTGC AAGATGGCCA CCCCATCGAT GCTGCCCCAG TGGGCGTACA TGCACATCGC
18361 CGGACAGGAC GCTTCGGAGT ACCTGAGTCC GGGTCTGGTG CAGTTCGCCC GCGCCACAGA
18421 CACCTACTTC AGTCTGGGGA ACAAGTTTAG GAACCCACG GTGGCGCCA CGCACGATGT
18481 GACCACCGAC CGCAGCCAGC GGCTGACGCT GCGCTTCGTG CCCGTGGACC GCGAGGACAA
18541 CACCTACTCG TACAAAGTGC GCTACACGCT GGCCGTGGG GACAACCGCG TGCTGGACAT
18601 GGCCAGCACC TACTTTGACA TCCGCGGCGT GCTGGATCGG GGCCCCAGCT TCAAACCCTA
18661 CTCCGGCACC GCCTACAACA GCCTGGCTCC CAAGGGAGCG CCCAACACCT CACAGTGGAT
18721 AACCAAAGAC AATGGAAC TGATAAGACATA CAGTTTTGGA AATGCTCCAG TCAGAGGATT
18781 GGACATTACA GAAGAGGGTC TCCAAATAGG ACCCGATGAG TCAGGGGGTG AAAGCAAGAA
18841 AATTTTTGCA GACAAAACCT ATCAGCCTGA ACCTCAGCTT GGAGATGAGG AATGGCATGA
18901 TACTATTGGA GCTGAAGACA AGTATGGAGG CAGAGCGCTT AAACCTGCCA CCAACATGAA
18961 ACCCTGCTAT GGGTCTTTCG CCAAGCCAAC TAATGCTAAG GGAGGTCAGG CTAAAAGCAG
19021 AACCAAGGAC GATGGCACTA CTGAGCCTGA TATTGACATG GCCTTCTTTG ACGATCGCAG
19081 TCAGCAAGCT AGTTTCAGTC CAGAACTTGT TTTGTATACT GAGAATGTCG ATCTGGACAC

Fig. 6K

SEQ ID NO:2

38/153

19141 CCCGGATACC CACATTATTT ACAAACCTGG CACTGATGAA ACAAGTTCTT CTTTCAACTT
19201 GGGTCAGCAG TCCATGCCCC ACAGACCCAA CTACATCGGC TTCAGAGACA ACTTTATCGG
19261 TCTCATGTAC TACAACAGTA CTGGCAATAT GGGTGTACTA GCTGGACAGG CCTCCCAGCT
19321 GAATGCTGTG GTGGACTTGC AGGACAGAAA CACTGAACTG TCCTACCAGC TCTTGCTTGA
19381 CTCTCTGGGT GACAGAACCA GGTATTTTCTAG TATGTGGAAC CAGGCGGTGG ACAGCTACGA
19441 CCCCATGTG CGCATTATTG AAAATCACGG TGTGGAGGAT GAACACCCA ACTATTGCTT
19501 CCCTTTGAAT GGTGTGGGCT TTACAGATAC ATTCCAGGGA ATTAAGGTTA AAACACCAA
19561 TAACGGAACA GCAAATGCTA CAGAGTGGGA ATCTGATACC TCTGTCAATA ATGCTAATGA
19621 GATTGCCAAG GGCAATCCTT TCGCCATGGA GATCAACATC CAGGCCAACC TGTGGCGGAA
19681 CTTCTCTTAC GCGAACGTGG CGCTGTACCT GCCCGACTCC TACAAGTACA CGCCGGCCAA
19741 CATCACGCTG CCCGCCAACA CCAACACCTA CGATTACATG AACGGCCGCG TGGTAGCGCC
19801 CTCGCTGGTG GACGCCCTACA TCAACATCGG GCGCGCTGG TCGCTGGACC CCATGGACAA
19861 CGTCAACCCC TTCAACCACC ACCGCAACGC GGGCCTGCGC TACCGCTCCA TGCTCCTGGG
19921 CAACGGGCGC TACGTGCCCT TCCACATCCA GGTGCCCCAA AAGTTTTTCG CCATCAAGAG
19981 CCTCCTGCTC CTGCCCCGGT CCTACACCTA CGAGTGAAC TTCCGCAAGG ACGTCAACAT
20041 GATCCTGCAG AGCTCCCTCG GCAACGACCT GCGCACGGAC GGGGCCTCCA TCGCCTTCAC
20101 CAGCATCAAC CTCTACGCCA CCTTCTTCCC CATGGCGCAC AACACCGCCT CCACGCTCGA
20161 GGCCATGCTG CGCAACGACA CCAACGACCA GTCCTTCAAC GACTACCTCT CGGCGGCCAA
20221 CATGCTCTAC CCCATCCCGG CCAACGCCAC CAACGTGCCC ATCTCCATCC CCTCGCGCAA
20281 CTGGGCGGCC TTCCGCGGCT GGTCTTTCAC GCGCCTCAAG ACCCGCGAGA CGCCCTCGCT
20341 CGGCTCCGGG TTCGACCCCT ACTTCGTCTA CTCGGGCTCC ATCCCTACC TCGACGGCAC
20401 CTTCTACCTC AACCACACCT TCAAGAAGGT CTCCATCACC TTCGACTCCT CCGTCAGCTG
20461 GCCCGGCAAC GACCGCCTCC TGACGCCCAA CGAGTTCGAA ATCAAGCGCA CCGTCGACGG
20521 AGAGGGGTAC AACGTGGCCC AGTGCAACAT GACCAAGGAC TGGTTCCTGG TTCAGATGCT
20581 GGCCCACTAC AACATCGGCT ACCAGGGCTT CTACGTGCCC GAGGGCTACA AGGACCGCAT
20641 GTACTCCTTC TTCCGCAACT TCCAGCCCAT GAGCCGCCAG GTCGTGGACG AGGTCAACTA
20701 CAAGGACTAC CAGGCCGTCA CCCTGGCCTA CCAGCACAAC AACTCGGGCT TCGTCGGCTA
20761 CCTCGCGCCC ACCATGCGCC AGGGACAGCC CTACCCCGCC AACTACCCCT ACCCGCTCAT
20821 CGGCAAGAGC GCCGTGCGCA GCGTCACCCA GAAAAAGTTC CTCTGCGACC GGGTCATGTG

Fig. 6L

SEQ ID NO:2

39/153

20881 GCGCATCCCC TTCTCCAGCA ACTTCATGTC CATGGGCGCG CTCACCGACC TCGGCCAGAA
20941 CATGCTCTAC GCCAACTCCG CCCACGCGCT AGACATGAAT TTCGAAGTCG ACCCCATGGA
21001 TGAGTCCACC CTTCTCTATG TTGTCTTCGA AGTCTTCGAC GTCGTCCGAG TGCACCAGCC
21061 CCACCGCGGC GTCATCGAGG CCGTCTACCT GCGCACGCCC TTCTCGGCCG GTAACGCCAC
21121 CACCTAAGCC CCGCTCTTGC TTCTTGCAAG ATGACGGCCT GTGCGGGCTC CGGCGAGCAG
21181 GAGCTCAGGG CCATCCTCCG CGACCTGGGC TCGGGGCCCT GCTTCCTGGG CACCTTCGAC
21241 AAGCGCTTCC CGGGATTTCAT GGCCCCGCAC AAGCTGGCCT GCGCCATCGT CAACACGGCC
21301 GGCCGCGAGA CCGGGGGCGA GCACTGGCTG GCCTTCGCCT GGAACCCGCG CTCCCACACC
21361 TGCTACCTCT TCGACCCCTT CGGGTTCTCG AACGAGCGCC TCAAGCAGAT CTACCAGTTC
21421 GAGTACGAGG GCCTGCTGCG CCGCAGCGCC CTGGCCACCG AGGACCGCTG CGTCACCTG
21481 GAAAAGTCCA CCCAGACCGT GCAGGGTCCG CGCTCGGCCG CCTGCGGGCT CTTCTGCTGC
21541 ATGTTCCCTGC ACGCCTTCGT GCACTGGCCC GACCGCCCCA TGGACAAGAA CCCCACCATG
21601 AACTTGCTGA CGGGGGTGCC CAACGGCATG CTCCAGTCGC CCCAGGTGGA ACCCACCTG
21661 CGCCGCAACC AGGAAGCGCT CTACCGCTTC CTCAACGCCC ACTCCGCCTA CTTTCGCTCC
21721 CACCGCGCGC GCATCGAGAA GGCCACCGCC TTCGACCGCA TGAATCAAGA CATGTAAACC
21781 GTGTGTGTAT GTGAATGCTT TATTCATAAT AAACAGCACA TGTTTATGCC ACCTTCTCTG
21841 AGGCTCTGAC TTTATTTAGA AATCGAAGGG GTTCTGCCGG CTCTCGGCAT GCGCGCGGG
21901 CAGGGATACG TTGCGGAAC TGTACTTGGG CAGCCACTTG AACTCGGGGA TCAGCAGCTT
21961 GGGCACGGGG AGGTGCGGGA ACGAGTCGCT CCACAGCTTG CGCGTGAGTT GCAGGGCGCC
22021 CAGCAGGTCG GCGCGGAGA TCTTGAAATC GCAGTTGGGA CCCGCGTTCT GCGCGCGAGA
22081 GTTGCGBTAC ACGGGGTTC AGCACTGGAA CACCATCAGG GCCGGGTGCT TCACGCTCGC
22141 CAGCACCGTC GCGTCGGTGA TGCCCTCCAC GTCCAGATCC TCGGCGTTGG CCATCCCGAA
22201 GGGGGTCATC TTGCAGGTCT GCCGCCCCAT GCTGGGCACG CAGCCGGGCT TGTGGTTGCA
22261 ATCGCAGTGC AGGGGGATCA GCATCATCTG GGCTTGCTCG GAGCTCATGC CCGGTACAT
22321 GGCTTTCATG AAAGCCTCCA GCTGGCGGAA GGCTTGCTGC GCCTTGCCGC CCTCGGTGAA
22381 GAAGACCCCG CAGGACTTGC TAGAGAACTG GTTGGTGGCG CAGCCCGCGT CGTGCACGCA
22441 GCAGCGCGCG TCGTTGTTGG CCAGCTGCAC CACGCTGCGC CCCAGCGGT TCTGGGTGAT
22501 CTTGGCCCGG TCGGGGTTCCT CCTTCAGCGC GCGCTGTCCG TTCTCGCTCG CCACATCCAT
22561 CTCGATCGTG TGCTCCTTCT GGATCATCAC GGTCCCGTGC AGGCACCGCA GCTTGCTCTC

Fig. 6M

ITR0048PV

SEQ ID NO:2

40/153

22621 GGCCCTCGGTG CACCCGTGCA GCCACAGCGC GCAGCCGGTG CTCTCCCAGT TCTTGTGGGC
22681 GATCTGGGAG TGCAGTGCA CGAAGCCCTG CAGGAAGCGG CCCATCATCG CGGTCAGGGT
22741 CTTGTTGCTG GTGAAGGTCA GCGGGATGCC GCGGTGCTCC TCGTTCACAT ACAGGTGGCA
22801 GATGCGGCGG TACACCTCGC CCTGCTCGGG CATCAGCTGG AAGGCGGACT TCAGGTGCT
22861 CTCCACGCGG TACCGGTCCA TCAGCAGCGT CATGACTTCC ATGCCCTTCT CCCAGGCCGA
22921 AACGATCGGC AGGCTCAGGG GGTTCCTTAC CGTTGTCATC TTAGTCGCCG CCGCCGAGGT
22981 CAGGGGGTCG TTCTCGTCCA GGTCTCAAA CACTCGCTTG CCGTCCTTCT CGATGATGCG
23041 CACGGGGGGG AAGCTGAAGC CCACGGCCGC CAGCTCCTCC TCGGCCTGCC TTTCGTCCTC
23101 GCTGTCCTGG CTGATGTCTT GCAAAGGCAC ATGCTTGGTC TTGCGGGGT TCTTTTTGGG
23161 CGGCAGAGGC GGCGGCGGAG ACGTGCTGGG CGAGCGCGAG TTCTCGCTCA CCACGACTAT
23221 TTCTTCTTCT TGGCCGTCGT CCGAGACCAC GCGGCGGTAG GCATGCCTCT TCTGGGGCAG
23281 AGGCGGAGGC GACGGGCTCT CGCGGTTCGG CCGGCGGCTG GCAGAGCCCC TTCCGCGTTC
23341 GGGGGTGCGC TCCTGGCGGC GCTGCTCTGA CTGACTTCCT CCGCGGCCGG CCATTGTGTT
23401 CTCCTAGGGA GCAACAACAA GCATGGAGAC TCAGCCATCG TCGCCAACAT CGCCATCTGC
23461 CCCC GCCGCC GACGAGAACC AGCAGAATGA AAGCTTAACC GCCCCGCCGC CCAGCCCCAC
23521 CTCCGACGCC GCGGCCCCAG ACATGCAAGA GATGGAGGAA TCCATCGAGA TTGACCTGGG
23581 CTACGTGACG CCCGCGGAGC ACGAGGAGGA GCTGGCAGCG CGCTTTTCAG CCCC GGAAGA
23641 GAACCACCAA GAGCAGCCAG AGCAGGAAGC AGAGAGCGAG CAGAACCAGG CTGGGCTCGA
23701 GCATGGCGAC TACCTGAGCG GGGCAGAGGA CGTGCTCATC AAGCATCTGA CCCGCCAATG
23761 CATCATCGTC AAGGACGCGC TGCTCGACCG CGCCGAGGTG CCCCTCAGCG TGGCGGAGCT
23821 CAGCCGCGCC TACGAGCGCA ACCTCTTCTC GCCGCGCGTG CCCCCAAGC GCCAGCCCAA
23881 CGGCACCTGC GAGCCCAACC CGCGCCTCAA CTTCTACCCG GTCTTCGCGG TGCCCGAGGC
23941 CCTGGCCACC TACCACCTCT TTTTCAAGAA CCAAAGGATC CCCGTCTCCT GCCGCGCCAA
24001 CCGCACCCGC GCCGACGCCC TGCTCAACCT GGGCCCCGGC GCCCGCCTAC CTGATATCAC
24061 CTCCTTGGA GAGGTTCCCA AGATCTTCGA GGTCTGGGC AGCGACGAGA CTCGGGCGCG
24121 GAACGCTCTG CAAGGAAGCG GAGAGGAACA TGAGCACCAC AGCGCCCTGG TGGAGTTGGA
24181 AGGCGACAAC GCGCGCCTGG CCGTGCTCAA GCGCACGGTC GAGCTGACCC ACTTCGCCTA
24241 CCCGGCGCTC AACCTGCCCC CCAAGGTCAT GAGCGCCGTC ATGGACCAGG TGCTCATCAA
24301 GCGCGCCTCG CCCATTGAGG ACATGCAGGA CCCC GAGAGC TCGGACGAGG GCAAGCCCGT

Fig. 6N

SEQ ID NO:2

41/153

24361 GGTCAAGCGAC GAGCAGCTGG CGCGCTGGCT GGGAGCGAGT AGCACCCCCC AGAGCCTGGA
24421 AGAGCGGCGC AAGCTCATGA TGGCCGTGGT CCTGGTGACC GTGGAGCTGG AGTGTCTGCG
24481 CCGCTTCTTC GCCGACGCAG AGACCCTGCG CAAGGTCGAG GAGAACCTGC ACTACCTCTT
24541 CAGGCACGGG TTCGTGCGCC AGGCCTGCAA GATCTCCAAC GTGGAGCTGA CCAACCTGGT
24601 CTCCTACATG GGCATCCTGC ACGAGAACCG CCTGGGGCAG AACGTGCTGC ACACCACCCT
24661 GCGCGGGGAG GCCCGCCGCG ACTACATCCG CGACTGCGTC TACCTGTACC TCTGCCACAC
24721 CTGGCAGACG GGCATGGGCG TGTGGCAGCA GTGCCTGGAG GAGCAGAACC TGAAAGAGCT
24781 CTGCAAGCTC CTGCAGAAGA ACCTCAAGGC CCTGTGGACC GGGTTCGACG AGCGCACCAC
24841 CGCCTCGGAC CTGGCCGACC TCATCTTCCC CGAGCGCCTG CGGCTGACGC TGCGCAACGG
24901 GCTGCCCCGAC TTTATGAGCC AAAGCATGTT GCAAACTTT CGCTCTTTCA TCCTCGAACG
24961 CTCCGGGATC CTGCCCCCA CCTGCTCCGC GCTGCCCTCG GACTTCGTGC CGCTGACCTT
25021 CCGCGAGTGC CCCCCGCGC TCTGGAGCCA CTGCTACTTG CTGCGCCTGG CCAACTACCT
25081 GGCCTACCAC TCGGACGTGA TCGAGGACGT CAGCGGCGAG GGTCTGCTGG AGTGCCACTG
25141 CCGCTGCAAC CTCTGCACGC CGCACCCTC CCTGGCCTGC AACCCCCAGC TGCTGAGCGA
25201 GACCCAGATC ATCGGCACCT TCGAGTTGCA AGGCCCCGGC GAGGAGGGCA AGGGGGGTCT
25261 GAAACTCACC CCGGGGCTGT GGACCTCGGC CTACTTGCGC AAGTTCGTGC CCGAGGACTA
25321 CCATCCCTTC GAGATCAGGT TCTACGAGGA CCAATCCCAG CCGCCCAAGG CCGAGCTGTC
25381 GGCCTGCGTC ATCACCAGG GGGCCATCCT GGCCCAATTG CAAGCCATCC AGAAATCCCG
25441 CCAAGAATTT CTGCTGAAAA AGGGCCACGG GGTCTACTTG GACCCCCAGA CCGGAGAGGA
25501 GCTCAACCCC AGCTTCCCCC AGGATGCCCC GAGGAAGCAG CAAGAAGCTG AAAGTGAGC
25561 TGCCGCCGCC GGAGGATTTG GAGGAAGACT GGGAGAGCAG TCAGGCAGAG GAGGAGATGG
25621 AAGACTGGGA CAGCACTCAG GCAGAGGAGG ACAGCCTGCA AGACAGTCTG GAGGAGGAAG
25681 ACGAGGTGGA GGAGGAGGAG GCAGAGGAAG AAGCAGCCGC CGCCAGACCG TCGTCTCCGG
25741 CGGAGAAAGC AAGCAGCACG GATACCATCT CCGCTCCGGG TCGGGGTCGC GCGGGCCGGG
25801 CCCACAGTAG GTGGGACGAG ACCGGGCGCT TCCCGAACCC CACCACCCAG ACCGGTAAGA
25861 AGGAGCGGCA GGGATACAAG TCCTGGCGGG GGCACAAAAA CGCCATCGTC TCCTGCTTGC
25921 AAGCCTGCGG GGGCAACATC TCCTTACCC GCGCTACCT GCTCTTCCAC CGCGGGGTGA
25981 ACTTCCCCCG CAACATCTTG CATTACTACC GTCACCTCCA CAGCCCCTAC TACTGTTTCC
26041 AAGAAGAGGC AGAAACCCAG CAGCAGCAGA AAACCAGCGA CAGCGGCAGC AGCTAGAAAA

Fig. 60

SEQ ID NO:2

42/153

26101 TCCACAGCGG CAGGTGGACT GAGGATCGCG GCGAACGAGC CGGCGCAGAC CCGGGAGCTG
26161 AGGAACCGGA TCTTTCCCAC CCTCTATGCC ATCTTCCAGC AGAGTCGGGG GCAGGAGCAG
26221 GAACTGAAAG TCAAGAACCG TTCTCTGCGC TCGCTCACCC GCAGTTGTCT GTATCACAAG
26281 AGCGAAGACC AACTTCAGCG CACTCTCGAG GACGCCGAGG CTCTCTTCAA CAAGTACTGC
26341 GCGCTCACTC TTAAAGAGTA GCCCGCGCCC GCCCACACAC GGAAAAAGGC GGGAAATTACG
26401 TCACCACCTG CGCCCTTCGC CCGACCATCA TCATGAGCAA AGAGATTCCC ACGCCTTACA
26461 TGTGGAGCTA CCAGCCCCAG ATGGGTCTGG CCGCCGGCGC CGCCAGGAC TACTCCACCC
26521 GCATGAACTG GCTCAGTGCC GGGCCCGCGA TGATCTCACG GGTGAATGAC ATCCGCGCCC
26581 ATCGAAACCA GATACTCCTA GAACAGTCAG CGATCACCGC CACGCCCCGC CATCACCTTA
26641 ATCCGCGTAA TTGGCCCGCC GCCCTGGTGT ACCAGGAAAT TCCCCAGCCC ACGACCGTAC
26701 TACTTCCGCG AGACGCCAG GCCGAAGTCC AGCTGACTAA CTCAGGTGTC CAGCTGGCCG
26761 GCGGCGCCGC CCTGTGTCGT CACCGCCCCG CTCAGGGTAT AAAGCGGCTG GTGATCCGAG
26821 GCAGAGGCAC ACAGCTCAAC GACGAGGTGG TGAGCTCTTC GCTGGGTCTG CGACCTGACG
26881 GAGTCTTCCA ACTCGCCGGA TCGGGGAGAT CTTCTTCAC GCCTCGTCAG GCCGTCTGA
26941 CTTTGAGAG TTCGTCCTCG CAGCCCCGCT CGGGCGGCAT CGGCACTCTC CAGTTCGTGG
27001 AGGAGTTCAC TCCCTCGGTC TACTTCAACC CCTTCTCCG CTCCCCGGC CACTACCCGG
27061 ACGAGTTCAT CCCGAACTTC GACGCCATCA GCGAGTCGGT GGACGGCTAC GATTGAATGT
27121 CCCATGGTGG CGCAGCTGAC CTAGCTCGGC TTCGACACCT GGACCACTGC CGCCGCTTCC
27181 GCTGCTTCGC TCGGGATCTC GCCGAGTTTG CCTACTTTGA GCTGCCCAG GAGCACCTC
27241 AGGGCCCGGC CCACGGAGTG CGGATCATCA TCGAAGGGG CCTCGACTCC CACCTGCTTC
27301 GGATCTTCAG CCAGCGACCG ATCCTGGTCG AGCGCGAGCA AGGACAGACC CGTCTGACCC
27361 TGTACTGCAT CTGCAACCAC CCCGGCCTGC ATGAAAGTCT TTGTTGTCTG CTGTGTACTG
27421 AGTATAATAA AAGCTGAGAT CAGCGACTAC TCCGGACTCG ATTGTGGTGT TCCTGCTATC
27481 AACCAGTCCC TGTTCTTCAC CGGGAACGAG ACCGAGCTCC AGCTCCAGTG TAAGCCCCAC
27541 AAGAAGTACC TCACCTGGCT GTTCCAGGGC TCTCCGATCG CCGTTGTCAA CCACTGCGAC
27601 AACGACGGAG TCCTGCTGAG CGGCCCTGCC AACCTTACTT TTTCCACCCG CAGAAGCAAG
27661 CTCCAGCTCT TCCAACCTT CCTCCCCGGG ACCTATCAGT GCGTCTCGGG ACCCTGCCAT
27721 CACACCTTCC ACCTGATCCC GAATACCACA GCGCCGCTCC CCGCTACTAA CAACCAAAC
27781 ACCCACCAAC GCCACCGTCG CGACCTTTC TCTGAATCTA ATACTACCAC CCACACCGGA

Fig. 6P

SEQ ID NO:2

43/153

27841 GGTGAGCTCC GAGGTCGACC AACCTCTGGG ATTTACTACG GCCCCTGGGA GGTGGTGGGG
27901 TTAATAGCGC TAGGCCTAGT TGTGGGTGGG CTTTGGGCTC TCTGCTACCT ATACCTCCCT
27961 TGCTGTTTCGT ACTTAGTGGT GCTGTGTTGC TGGTTTAAGA AATGGGGAAG ATCACCCTAG
28021 TGAGCTGCGG TGTGCTGGTG GCGGTGTTGC TTTCGATTGT GGGACTGGGC GGC GCGGCTG
28081 TAGTGAAGGA GGAGAAGGCC GATCCCTGCT TGCATTTCAA TCCCACAAA TGCCAGCTGA
28141 GTTTTCAGCC CGATGGCAAT CCGTGCACGG TGCTGATCAA GTGCGGATGG GAATGTGAGA
28201 ACGTGAGAAT CGAGTACAAT AACAAGACTC GGAACAATAC TCTCGCGTCC GTGTGGCAAC
28261 CCGGGGACCC CGAGTGGTAC ACCGTCTCTG TCCCCGGTGC TGACGGCTCC CCGCGCACCG
28321 TGAATAATAC TTTCATTTTT GCGCACATGT GCGACACGGT CATGTGGATG AGCAAGCAGT
28381 ACGATATGTG GCGGGGACG AAGGAGAACA TCGTGGTCTT CTCCATCGCT TACAGCGTGT
28441 GCACGGCGCT AATCACCCT ATCGTGTGCC TGAGCATTCA CATGCTCATC GCTATTCGCC
28501 CCAGAAATAA TGCCGAAAAA GAGAAACAGC CATAACACGT TTTTTCACAC ACCTTTTTCA
28561 GACCATGGCC TCTGTTACTG CCCTAATTAT TTTTTTGGGT CTCGTGGGCA CTAGCAGCAC
28621 TTTTCAGCAT ATAAACAAAA CTGTTTATGC TGGTCTAAT TCTGTATTAC CTGGGCATCA
28681 ATCACACCAG AAAGTTTCAT GGTACTGGTA TGATAAAAAT AACACGCCAG TCACACTCTG
28741 CAAGGGTCAT CAAACACCCA TAAACCGTAG TGGAATTTTT TTAAATGTA ATCATAATAA
28801 TATTACACTA CTTTCAATTA CAAAGCACTA TTCTGGTACT TACTATGGAA CCAATTTTAA
28861 CATAAAACAG GACACTTACT ATAGTGTCAC AGTATTGGAT CCAACTACTC CTAGAACAAC
28921 TACAAAACCC ACAACTACTA AGAGGCACAC TAAACCTAAA ACTACCAAGA AAACCACTGT
28981 CAAAACAACA ACTAGGACCA CCACAACCTAC AGAGGCTACC ACCAGCACAA CACTTGCTGC
29041 AACTACACAC ACACACACTG AGCTAACCTT ACAGACCACT AATGATTTGA TAGCCCTGTT
29101 GCAAAAGGGG GATAACAGCA CCACTTCCGA TGAGGAAATA CCCAAATCCA TGATTGGCAT
29161 TATTGTTGCT GTAGTGGTGT GCATGTTGAT CATCGCCTTG TGCATGGTGT ACTATGCCTT
29221 CTGCTACAGA AAGCACAGAC TGAACGACAA GCTGGAACAC TTACTAAGTG TTGAATTTTA
29281 ATTTTTTTAGA ACCATGAAGA TCCTAGGCCT TTTAGTTTTT TCTATCATTA CCTCTGCTCT
29341 TTGTGAATCA GTGAATAAAG ATGTTACTAT TACCACTGGT TCTAATTATA CACTGAAAGG
29401 GCCACCCTCA GGTATGCTTT CGTGGTATTG CTATTTTGGG ACTGACACTG ATCAAACCTGA
29461 ATTATGCAAT TTTCAAAAAG GCAAAACCTC AAACCTCTAAA ATCTCTAATT ATCAATGCAA
29521 TGGCACTGAT CTGATACTAC TCAATGTCAC GAAAGCATAT GGTGGCAGTT ATTCTTGCCC

Fig. 6Q

SEQ ID NO:2

44/153

29581 TGGACAAAAC ACTGAAGAAA TGATTTTTTA CAAAGTGGAA GTGGTTGATC CCACTACTCC
29641 ACCCACCACC ACAACTACTC ACACCACACA CACAGAACAA ACCACAGCAG AGGAGGCAGC
29701 AAAGTTAGCC TTGCAGGTCC AAGACAGTTC ATTTGTTGGC ATTACCCCTA CACCTGATCA
29761 GCGGTGTCCG GGGCTGCTAG TCAGCGGCAT TGTCGGTGTG CTTTCGGGAT TAGCAGTCAT
29821 AATCATCTGC ATGTTCAATTT TTGCTTGCTG CTATAGAAGG CTTTACCGAC AAAAATCAGA
29881 CCCACTGCTG AACCTCTATG TTTAATTTTT TCCAGAGCCA TGAAGGCAGT TAGCACTCTA
29941 GTTTTTTGTT CTTTGATTGG CATGTTTTTT AGTGCTGGGT TTTTGAAAAA TCTTACCATT
30001 TATGAAGGTG AGAATGCCAC TCTAGTGGGC ATCAGTGGTC AAAATGTCAG CTGGCTAAAA
30061 TACCATCTAG ATGGGTGGAA AGACATTGTC GATTGGAATG TCACTGTGTA TACATGTAAT
30121 GGAGTTAACC TCACCATTAC TAATGCCACC CAAGATCAGA ATGGTAGGTT TAAGGGTCAG
30181 AGTTTCACTA GAAATAATGG GTATGAATCC CATAACATGT TTATCTATGA CGTCACTGTC
30241 ATCAGAAATG AGACCGCCAC CACCACACAG ATGCCCACTA CACACAGTTC TACCACTACT
30301 ACCAAGCAAA CCACACAGAC AACCACTTTT TATACATCAA CTCAGCATAT GACCACCACT
30361 ACAGCAGCAA AGCCAAGTAG CGCAGCGCCT CAGCCACAGG CTTTGGCTTT GAAAGCTGCA
30421 CAACCTAGTA CAACTACTAA GACCAATGAG CAGACTACTG ATTTTTTGTC CACTGTGCGAG
30481 AGCCACACCA CAGCTACCTC CAGTGCCCTT TCTAGCACCG CCAATCTCTC CTCGCTTTCC
30541 TCTACACCAA TCAGTCCCGC TACTACTCCT AGCCCCGCTC CTCTTCCCAC TCCCCTGAAG
30601 CAAACAGACG GCGGCATGCA ATGGCAGATC ACCCTGCTCA TTGTGATCGG GTTGGTCATC
30661 CTGGCCGTGT TGCTCTACTA CATCTTCTGC CGCCGCATTC CCAACGCGCA CCGCAAGCCG
30721 GTCTACAAGC CCATCGTTGT CGGGCAGCCG GAGCCGCTTC AGGTGGAAGG GGGTCTAAGG
30781 AATCTTCTCT TCTCTTTTAC AGTATGGTGA TTGAACTATG ATTCCTAGAC AATTCTTGAT
30841 CACTATTCTT ATCTGCCTCC TCCAAGTCTG TGCCACCCTC GCTCTGGTGG CCAACGCCAG
30901 TCCAGACTGT ATTGGGCCCT TCGCCTCCTA CGTGCTCTTT GCCTTCATCA CCTGCATCTG
30961 CTGTTGTAGC ATAGTCTGCC TGCTTATCAC CTTCTTCCAG TTCATTGACT GGATCTTTGT
31021 GCGCATCGCC TACCTGCGCC ACCACCCCCA GTACCGCGAC CAGCGAGTGG CGCGACTGCT
31081 CAGGCTCCTC TGATAAGCAT GCGGGCTCTG CTACTTCTCG CGCTTCTGCT GTTAGTGCTC
31141 CCCCCTCCCG TCGACCCCCG GTCCCCCGAG GAGGTCCGCA AATGCAAATT CCAAGAACCC
31201 TGGAAATTCC TCAAATGCTA CCGCCAAAAA TCAGACATGC ATCCCAGCTG GATCATGATC
31261 ATTGGGATCG TGAACATTCT GGCCTGCACC CTCATCTCCT TTGTGATTTA CCCCTGCTTT

Fig. 6R

SEQ ID NO:2

45/153

31321 GACTTTGGTT GGAAGCTCGCC AGAGGCACTC TATCTCCCGC CTGAGCCTGA CACACCACCA
31381 CAGCAGCAAC CTCAGGCACA CGCACTACCA CCACCACAGC CTAGGCCACA ATACATGCCC
31441 ATATTAGACT ATGAGGCCGA GCCACAGCGA CCCATGCTCC CCGCTATTAG TTAATTCAAT
31501 CTAACCGGCG GAGATGACTG ACCCACTGGC CAACAACAAC GTCAACGACC TTCTCCTGGA
31561 CATGGACGGC CGCGCCTCGG AGCAGCGACT CGCCCAACTC CGCATCCGCC AGCAGCAGGA
31621 GAGAGCCGTC AAGGAGCTGC AGGATGCGGT GGCCATCCAC CAGTGCAAGA AAGGCATCTT
31681 CTGCCTGGTG AAGCAGGCCA AGATCTCCTA CGAGGTCACC CAGACCGACC ATCGCCTCTC
31741 CTACGAGCTC CTGCAGCAGC GCCAGAAGTT CACCTGCCTG GTCGGAGTCA ACCCCATCGT
31801 CATCACCCAG CAGTCGGGCG ATACCAAGGG GTGCATCCAC TGCTCCTGCG ACTCCCCCGA
31861 GTGCGTTTAC ACCATGATCA AGACCTCTG CGGCCTCCGC GACCTCCTCC CCATGAACATA
31921 ATCACCCCTT TATCCAGTGA AATAAAGATC ATATTGATGA TGATTTAAAT AAAAAATAA
31981 TCATTTGATT TGAAATAAAG ATACAATCAT ATTGATGATT TGAGTTTAAC AAAAATAAAG
32041 AATCACTTAC TTGAAATCTG ATACCAGGTC TCTGTCCATG TTTTCTGCCA ACACCACCTC
32101 ACTCCCCTCT TCCCAGCTCT GGTACTGCAG GCCCCGGCGG GCTGCAAACT TCCTCCACAC
32161 GCTGAAGGGG ATGTCAAATT CCTCCTGTCC CTCAATCTTC ATTTTCTCTT CTATCAGATG
32221 TCCAAAAAGC GCGCGCGGGT GGATGATGAC TTCGACCCCG TGTACCCCTA CGATGCAGAC
32281 AACGCACCGA CTGTGCCCTT CATCAACCCT CCCTTCGTCT CTTTCAAGATG ATTCCAAGAA
32341 AAGCCCCTGG GGGTGTGTGTC CCTGCGACTG GCCGATCCCG TCACCACCAA GAACGGGGCT
32401 GTCACCCTCA AGCTGGGGGA GGGGGTGGAC CTCGACGACT CGGGAAAACT CATCTCCAAA
32461 AATGCCACCA AGGCCACTGC CCCTCTCAGT ATTTCCAACA ACACCATTC CCTTAACATG
32521 GATACCCCTC TTTACAACAA CAATGGAAAG CTAGGTATGA AGGTAACCGC ACCATTAAAG
32581 ATATTAGACA CAGATCTACT AAAAACAATT GTTGTGTGCTT ATGGGCAGGG ATTAGGAACA
32641 AACACCAATG GTGCTCTTGT TGCCCAACTA GCATACCCAC TTGTTTTTAA TACCGCTAGC
32701 AAAATTGCCC TTAATTTAGG CAATGGACCA TTAAAAGTGG ATGCAAATAG ACTGAACATT
32761 AATTGCAAAA GAGGTATCTA TGCTACTACC ACAAAGATG CACTGGAGAT TAATATCAGT
32821 TGGGCAAATG CTATGACATT TATAGGAAAT GCCATTGGTG TCAATATTGA CACAAAAAAA
32881 GGCCTACAGT TCGGCACTTC AAGCACTGAA ACAGATGTTA AAAATGCTTT TCCACTCCAA
32941 GTAAACTTG GAGCTGGTCT TACATTTGAC AGCACAGGTG CCATTGTTGC TTGGAACAAA
33001 GAAGATGACA AACTTACACT GTGGACCACA GCCGATCCAT CTCCAAACTG TCACATATAT

Fig. 6S

SEQ ID NO:2

46/153

33061 TCTGCAAAGG ATGCTAAGCT TACACTCTGC TTGACAAAGT GTGGTAGTCA GATACTGGGC
33121 ACTGTTTCTC TCATAGCTGT TGATACTGGT AGCTTAAATC CAATAACAGG AAAAGTAACC
33181 ACTGCTCTTG TTTCAC TTAA ATTCGATGCC AATGGAGTTT TGCAAGCCAG TTCAACACTA
33241 GATAAAGAAT ATTGGAATTT CAGAAAAGGA GATGTGACAC CTGCTGACCC CTACACTAAT
33301 GCTATAGGCT TTATGCCCAA CCTTAATGCA TACCCAAAAA ACACAAACGC AGCTGCAAAA
33361 AGTCACATTG TTGGAAAAGT ATACCTACAT GGGGATGAAA GCAAGCCACT AGACTTGATA
33421 ATTACATTTA ATGAAACCAG TGATGAATCC TGTACTTATT GCATTAAC TT TCAGTGGCAG
33481 TGGGGAAGT ACCAATATAA AGATGAAACA CTTGCAGTCA GTTCATTAC CTTCTCATAC
33541 ATTGCTAAAG AATAACATCC ACCCTGCATG CCAACCCATT TCCCTCTATC TATACATGGA
33601 AACTCTGAA GCAGAAAAAA TAAAGTTCAA GTGTTTTATT GATTCAACAG TTTTACAGA
33661 ATTCGAGTAG TTATTTTCCC TCCACCCTCC CAACTCATGG AATACACCAT CCTCTCCCCA
33721 CGCACAGCCT TAAACATCTG AATGCCATTG GTAATGGACA TGGTTTGGC CTCCACATTC
33781 CACACAGTTT CAGAGCGAGC CAGTCTCGGG TCGGTCAGGG AGATGAAACC CTCCGGGCAC
33841 TCCTGCATCT GCACCTCACA GTTCAACAGC TGAGGGCTGT CCTCGGTGGT CGGGATCACA
33901 GTTATCTGGA AGAAGAGCGA TGAGAGTCAT AATCCGCGAA CGGGATCGGG CGGTTGTGGC
33961 GCATCAGGCC CCGCAGCAGT CGCTGTCTGC GCCGCTCCGT CAAGCTGCTG CTCAAGGGGT
34021 CCGGGTCCAG GGA CTCCCCG CGCATGATGC CGATGGCCCT GAGCATCAGT CGCCTGGTGC
34081 GCGGGGCGCA GCAGCGGATG CGGATCTCAC TCAGGTCGGA ACAGTACGTG CAGCACAGCA
34141 CTACCAAGTT GTTCAACAGT CCATAGTTCA ACGTGCTCCA GCCAAAAC TC ATCTGTGGAA
34201 CTATGCTGCC CACATGTCCA TCGTACCAGA TCCTGATGTA AATCAGGTGG CGCCCCCTCC
34261 AGAACACACT GCCCATGTAC ATGATCTCCT TGGGCATGTG CAGGTTACC ACCTCCCGGT
34321 ACCACATCAC CCGCTGGTTG AACATGCAGC CCCGGATGAT CCTGCGGAAC CACAGGGCCA
34381 GCACCGCCCC GCCCGCCATG CAGCGCAGGG ACCCGGGTC CTGGCAATGG CAGTGGATGA
34441 TCCACCGCTC GTACCCGTGG ATCATCTGGG AGCTGAACAA GTCTATGTTG GCACAGCACA
34501 GGCACACGCT CATGCATCTC TTCAGCACTC TCAGCTCCTC GGGGGTCAAA ACCATATCCC
34561 AGGGTACGGG GAACTCTTGC AGGACAGCGA ACCCGCAGA ACAGGGCAAA CCTCGCACAG
34621 AACTTACATT GTGCATGGAC AGGGTATCGC AATCAGGCAG CACCGGGTGA TCCTCCACCA
34681 GGAAGCGCG GGTCTCGATT TCCTCACAGC GTGGTAAGGG GCGCGGTCGA TACGGGTGAT
34741 GGCGGGACGC GGCTGATCGT GTTCGCGATC GTGTCATGAT GCAGTTGCTT TCGGACATTT

Fig. 6T

SEQ ID NO:2

47/153

34801 TCGTACTTGC TATAGCAGAA CCTGGTCCGG GCGCTGCACA CCGATCGCCG GCGGCGGTCT
34861 GCGCGCTTGG AACGCTCCGT GTTGAAATTG TAAAACAGCC ACTCTCTCAG ACCGTGCAGC
34921 AGATCTAGGG CCTCAGGAGT GATGAAGATC CCATCATGCC TGATGGCTCT GATCACATCG
34981 ACCACCGTGG AATGGGCCAG ACCCAGCCAG ATGATGCAAT TTTGTTGGGT TTCGGTGACG
35041 GCGGGGGAGG GAAGAACAGG AAGAACCATG ATTAACTTTA ATCCAAACGG TCTCGGAGCA
35101 CTTCAAAATG AAGGTCGCGG AGATGGCACC TCTCGCCCCC GCTGTGTTGG TGGAAAATAA
35161 CAGCCAGGTC AAAGGTGATA CGGTTCTCGA GATGTTCCAC GGTGGCTTCC AGCAAAGCCT
35221 CCACGCGCAC ATCCAGAAAC AAGACAATAG CGAAAGCGGG AGGGTTCTCT AATTCCTCAA
35281 TCATCATGTT ACACTCCTGC ACCATCCCCA GATAATTTTC ATTTTCCAG CCTTGAATGA
35341 TTCGAACTAG TTCCTGAGGT AAATCCAAGC CAGCCATGAT AAAGAGCTCG CGCAGAGCGC
35401 CCTCCACCGG CATTCTTAAG CACACCCTCA TAATTCCAAG ATATTCTGCT CCTGGTTCAC
35461 CTGCAGCAGA TTGACAAGCG GAATATCAAA CTCTCTGCCG CGATCCCTAA GCTCCTCCCT
35521 CAGCAATAAC TGTAAGTACT CTCTCATATC CTCTCCGAAA TTTTGTAGCA TAGGACCGCC
35581 AGGAATAAGA TTAGGGCAAG CCACAGTACA GATAAACCGA AGTCCTCCCC AGTGAGCATT
35641 GCCAAATGCA AGACTGCTAT AAGCATGCTG GCTAGACCCG GTGATATCTT CCAGATAATT
35701 GGACAGAAAA TCGCCCAGGC AATTTTAAAG AAAATCAACA AAAGAAAAAT CCTCCAGGTG
35761 CACGTTTAGA GCCTCGGGAA CAACGATGGA GTAAATGCAA GCGGTGCGTT CCAGCATGGT
35821 TAGTTAGCTG ATCTGTAGAA AAAACAAAAA TGAACATTAA ACCATGCTAG CCTGGCGAAC
35881 AGGTGGGTAA ATCGTTCTTT CCAGCACCAG GCAGGCCACG GGGTCTCCGG CGCGACCCTC
35941 GTAAAAATTG TCGCTATGAT TGAAAACCAT CACAGAGAGA CGTTCCCGGT GGCCGGCGTG
36001 AATGATTGCA CAAGACGAAT ACACCCCCGG AACATTGGCG TCCGCGAGTG AAAAAAGCG
36061 CCCGAGGAAG CAATAAGGCA CTACAATGCT CAGTCTCAAG TCCAGCAAAG CGATGCCATG
36121 CGGATGAAGC ACAAATTTCT CAGGTGCGTA CAAAATGTAA TTACTCCCTT CCTGCACAGG
36181 CAGCAAAGCC CCCGATCCCT CCAGGTACAC ATACAAAGCC TCAGCGTCCA TAGCTTACCG
36241 AGCAGCAGCG GCACACAACA GCGCAAAAAG TCAGAGAAAG GCTGAGAGCT CTAACCTGTC
36301 CACCCGCTCT CTGCTCAATA TATAGCCCAG ATCTACACTG ACGTAAAGGC CAAAGTCTAA
36361 AAATACCCGC CAAATAATCA CACACGCCCC GCACACGCCC AGAAACCGGT GACACACTCA
36421 GAAAAATACG CGCACTTCCT CAAACGCCCC AACTGCCGTC ATTTCCGGGT TCCCACGCTA
36481 CGTCATCAAA ATTCAACTTT CAAATTCCGT CGACCGTTAA AAACGTCACC CGCCCCGCCC

Fig. 6U

ITR0048PV

SEQ ID NO:2

48/153

36541 CTAACGGTCG CCGCTCCCGC AGCCAATCAG CGCCCCGCAT CCCCAAATTC AAACGGCTCA
36601 TTTGCATATT AACGCGCACC AAAAGTTTGA GGTATATTAT TGATGATG

Fig. 6V

SEQ ID NO:3

49/153

```

1  CATCATCAAT AATATACCTC AAACCTTTTGG TGC GCGTTAA TATGCAAATG AGCTGTTTGA
61  ATTTGGGGAG GGAGGAAGGT GATTGGCTGC GGGAGCGGCG ACCGTTAGGG GCGGGGCGGG
121 TGACGTTTTG ATGACGTGGC TATGAGCGCG AGCCGGTTTG CAAGTTCTCG TGGGAAAAGT
181 GACGTCAAAC GAGGTGTGGT TTGAACACGG AAATACTCAA TTTTCCCGCG CTCTCTGACA
241 GGAATGAGG TGTCTCTGGG CGGATGCAAG TGAAAACGGG CCATTTTCGC GCGAAAAC TG
301 AATGAGGAAG TGAAAATCTG AGTAATTTTCG CGTTTATGGC AGGGAGGAGT ATTTGCCGAG
361 GGCCGAGTAG ACTTTGACCG ATTACGTGGG GGTTCGATT ACCGTATTTT TCACCTAAAT
421 TTCCGCGTAC GGTGTCAAAG TCCGGTGT TTACGTAGGC GTCAGCTGAT CGCCAGGGTA
481 TTAAACCTG CGCTCTCTAG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGTTTCT
541 CCTCCGCGCC GCGAGTCAGA TCTACACTTT GAAAGATGAG GCACCTGAGA GACCTGCCCG
601 GTAATGTTTT CCTGGCTACT GGAACGAGA TTCTGGAATT GGTGGTGGAC GGCATGATGG
661 GTGACGACCC TCCAGAGCCC CTACCCCAT TTGAGGCGCC TTCGCTGTAC GATTGTATG
721 ATCTGGAGGT GGATGTGCCC GAGAGCGACC CTAACGAGGA GGCGGTGAAT GATTGTGTTA
781 GCGATGCCGC GCTGCTGGCT GCCGAGCAGG CTAATACGGA CTCTGGCTCA GACAGCGATT
841 CCTCTCTCCA TACCCCGAGA CCCGGCAGAG GTGAGAAAAA GATCCCGGAG CTAAAGGGG
901 AAGAGCTCGA CTGCGCTGC TATGAGGAAT GCTTGCTCC GAGCGATGAT GAGGAGGACG
961 AGGAGGCGAT TCGAGCTGCG GTGAACGAG GAGTGA AAC TGC GGGCGAG AGCTTTAGCC
1021 TGGACTGTCC TACTCTGCCC GGACACGGCT GTAAGTCTTG TGAATTTTCA TGCATGAATA
1081 CTGGAGATAA GAATGTGATG TGTGCCCTGT GCTATATGAG AGCTTACAAC CATTTGTTT
1141 ACAGTAAGTG TGATTAAC TTGTTGGGAA GGCAGAGGGT GACTGGGTGC TGACTGGTTT
1201 ATTTATGTAT ATGTTTTTTT ATGTGTAGGT CCCGTCTCTG ACGTAGATGA GACCCCACT
1261 TCAGAGTGCA TTTCATCACC CCCAGAAATT GGCGAGGAAC CGCCCGAAGA TATTATTCAT
1321 AGACCAGTTG CAGTGAGAGT CACCGGGCGG AGAGCAGCTG TGGAGAGTTT GGATGACTTG
1381 CTACAGGGTG GGGATGAACC TTGGACTTG TGTACCCGGA AACGCCCCAG GCATAAGTG
1441 CCACACATGT GTGTTTACTT AAGGTGATGT CAGTATTTAT AGGCTGTGGA GTGCAATAAA
1501 ATCCGTGTTG ACTTTAAGTG CGTGTTTTAT GACTCAGGGG TGGGGACTGT GGGTATATAA
1561 GCAGGTGCAG ACCTGTGTGG TCAGTTCAGA GCAGGACTCA TGGAGATCTG GACTGTCTTG
1621 GAAGACTTTC ACCAGACTAG ACAGTTGCTA GAGAACTCAT CGGAGGGAGT CTCTTACCTG
1681 TGGAGATTCT GCTTCGGTGG GCCTCTAGCT AAGCTAGTCT ATAGGGCCAA ACAGGATTAT
1741 AAGGAACAAT TTGAGGATAT TTTGAGAGAG TGTCTTGGA TTTTGGACTC TCTCAACTTG
1801 GGCCATCAGT CTCACTTAA CCAGAGATT CTGAGAGCCC TTGACTTTTC TACTCTGGC
1861 AGAACTACCG CCGCGGTAGC CTTTTTTGCC TTTATTCTTG ACAAATGGAG TCAAGAAACC
1921 CATTTTCAGCA GGGATTACCG TCTGGACTGC TTAGCAGTAG CTTTGTGGAG AACATGGAGG
1981 TGCCAGCGCC TGAATGCAAT CTCCGGCTAC TTGCCAGTAC AGCCGGTAGA CACGCTGAGG
2041 ATCCTGAGTC TCCAGTCACC CCAGGAACAC CAACGCCGCC AGCAGCCGCA GCAGGAGCAG
2101 CAGCAAGAGG AGGACCGAGA AGAGAACCCG AGAGCCGGTC TGGACCTCC GGTGGCGGAG
2161 GAGGAGGAGT AGCTGACTTG TTTCCCGAGC TGC GCGGGT GCTGACTAGG TCTTCCAGTG
2221 GACGGGAGAG GGGGATTAAG CGGGAGAGC ATGAGGAGAC TAGCCACAGA ACTGAAGTGA
2281 CTGTCACTCT GATGAGCCGC AGGCGCCAG AATCGGTGTG GTGGCATGAG GTGCACTGCG
2341 AGGGGATAGA TGAGGTCTCG GTGATGCATG AGAAATATTC CCTAGAACAA GTCAAGACTT
2401 GTTGGTTGGA GCCCGAGGAT GATTGGGAGG TAGCCATCAG GAATTATGCC AAGCTGGCTC
2461 TGAAGCCAGA CAAGAAGTAC AAGATTACCA AACTGATTAA TATCAGAAAT TCCTGCTACA
2521 TTTCAGGGAA TGGGGCCGAG GTGGAGATCA GTACCCAGGA GAGGGTGGCC TTCAGATGTT
2581 GTATGATGAA TATGTACCCG GGGGTGGTGG GCATGGAGGG AGTCACCTTT ATGAACACGA
2641 GGTTACGGGG TGATGGGTAT AATGGGGTGG TCTTTATGGC CAACACCAAG CTGACAGTGC
2701 ACGGATGCTC CTTCTTTGGC TTCAATAACA TGTGCATCGA GGCCTGGGGC AGTGTTCAG
2761 TGAGGGGATG CAGCTTTTCA GCCAACTGGA TGGGGGTCTG GGCAGAAACC AAGAGCAAGG
2821 TGTCAGTGAA GAAATGCCTG TTCGAGAGGT GCCACCTGGG GGTGATGAGC GAGGGCGAAG
2881 CCAAAGTCAA AACTGCGCC TCTACCGAGA CGGGCTGCTT TGTGCTGATC AAGGGCAATG
2941 CCAAAGTCAA GCATAACATG ATCTGTGGGG CCTCGGATGA GCGCGGCTAC CAGATGCTGA
3001 CCGCGCCGG TGGGAACAGC CATATGCTGG CCACCGTGCA TGTGGCCTCG CACCCCGCA
3061 AGACATGGCC CGAGTTTCGAG CACAACGTCA TGACCCGCTG CAATGTGCAC CTGGGCTCCC
3121 GCCGAGGCAT GTTCATGCC TACCAGTGCA ACATGCAATT TGTGAAGGTG CTGCTGGAGC
3181 CCGATGCCAT GTCCAGAGTG AGCCTGACGG GGGTGTGTTGA CATGAATGTG GAGCTGTGGA
3241 AAATTCTGAG ATATGATGAA TCCAAGACCA GGTGCCGGGC CTGCGAATGC GGAGGCAAGC
3301 ACGCCAGGCT TCAGCCCGTG TGTGTGGAGG TGACGGAGGA CCTGCGACCC GATCATTTGG
3361 TGTGTCTCTG CAACGGGACG GAGTTCGGCT CCAGCGGGGA AGAATCTGAC TAGAGTGAGT
3421 AGTGTGTTGG GCTGGGTGTG AGCCTGCATG AGGGGCAGAA TGACTAAAAT CTGTGTTTTT
3481 CTGTGTGTTG CAGCAGCATG AGCGGAAGCG CCTCCTTTGA GGGAGGGGTA TTCAGCCCTT
3541 ATCTGACGGG GCGTCTCCCC TCCTGGGCGG GAGTGTGTCA GAATGTTATG GNATCCACGG
3601 TGGACGGCCG GCGCGTGCAG CCCGCGAACT CTTCAACCCT GACCTACGCG ACCCTGAGCT

```

Fig. 7A

3661 CCTCGTCCGT GGACGCAGCT GCCGCCGCGAG CTGCTGCTTC CGCCGCCAGC GCCGTGCGCG
3721 GAATGGCCCT GGGCGCCGGC TACTACAGCT CTCTGGTGGC CAACTCGAGT TCCACCAATA
3781 ATCCCGCCAG CCTGAACGAG GAGAAGCTGC TGCTGCTGAT GGCCAGCTC GAGGCCCTGA
3841 CCCAGCGCCT GGGCGAGCTG ACCCAGCAGG TGGCTCAGCT GCAGGCGGAG ACGCGGGCCG
3901 CGGTTGCCAC GGTGAAAACC AAATAAAAAA TGAATCAATA AATAAACGGA GACGGTTGTT
3961 GATTTTAACA CAGAGTCTTG AATCTTTAT TGAATTTTCG CGCGCGGTAG GCCCTGGACC
4021 ACCGGTCTCG ATCATTGAGC ACCCGGTGGA TCTTTTCCAG GACCCGGTAG AGGTGGGCTT
4081 GGATGGTGAG GTACATGGGC ATGAGCCCGT CCCGGGGGTG GAGGTAGCTC CATTCAGGG
4141 CCTCGTGCTC GGGGATGGTG TTGTAAATCA CCCAGTCATA GCAGGGGCGC AGGGCGTGGT
4201 GCTGCACGAT GTCCTTGAGG AGGAGACTGA TGGCCACGGG CAGCCCTTG GTGTAGGTGT
4261 TGACGAACCT GTTGAGCTGG GAGGGATGCA TGCGGGGGGA GATGAGATGC ATCTTGGCTT
4321 GGATCTTGAG ATTGGCGATG TTCCCGCCCA GATCCCGCCG GGGGTTTCATG TTGTGCAGGA
4381 CCACCAGCAC GGTGTATCCG GTGCACTTGG GGAATTTGTC ATGCAACTTG GAAGGGAAGG
4441 CGTGAAAGAA TTTGGAGACG CCCTTGAGC CGCCAGGTT TTCCATGCAC TCATCCATGA
4501 TGATGGCTGAT GGGCCCGTGG GCGGCGGCCT GGGCAAAGAC GTTTCGGGGG TCGGACACAT
4561 CGTAGTTGTG GTCTGGGTG AGCTCGTCAT AGGCCATTTT AATGAATTTG GGGCGGAGGG
4621 TGCCCCACTG GGGGACGAAG GTGCCCTCGA TCCCGGGGGC GTAGTTGCCC TCGCAGATGT
4681 GCATCTCCCA GGCCTTGAGC TCGGAGGGGG GGATCATGTC CACCTGCGGG GCGATGAAAA
4741 AAACGGTTTT CCGGCGCGGG GAGATGAGCT GGGCCGAAAG CAGGTTCCGG AGCAGCTGGG
4801 ACTTGCCGCA ACCGGTGGGG CCGTAGATGA CCCCAGTAC CGGCTGCAGG TGGTAGTTGA
4861 GGGAGAGACA GCTGCCGTCC TCGCGGAGGA GGGGGGCCAC CTCGTTTCATC ATCTCGCGCA
4921 CATGCATGTT CTCGCGCAG AGTTCCGCCA GGAGGCGCTC GCGCATGGGC ATTTTGGAGA
4981 CTTGCAGCGA GCGCAAGTTT TTCAGCGGCT TGAGTCCGTC GGCCATGGGC ATTTTGGAGA
5041 GGGTCTGTTG CAAGAGTTCC AGACGGTCCC AGAGCTCGGT GATGTGCTCT AGGGCATCTC
5101 GATCCAGCAG ACCTCCTCGT TTCGCGGGT GGGGCGACTG CCGGAGTAGG GCACCAGGCG
5161 ATGGGCGTCC AGCGAGGCCA GGGTCCGGTC CTTCCAGGGC CGCAGGGTCC GCGTCAGCGT
5221 GGTCCTCCGT ACGGTGAAGG GGTGCGCGCC GGGCTGGCG CTTGCGAGGG TCGCTTCAG
5281 GCTCATCCGG CTGGTCGAGA ACCGCTCCCG GTCCGCGCCC TGCGCGTCGG CCAGGTAGCA
5341 ATTGAGCATG AGTTCGTAGT TGAGCGCCTC GGCCGCGTGG CCTTGGCGC GGAGCTTACC
5401 TTTGGAAGTG TGTCCGAGAG CGGGACGAG GAGGGACTTG AGGGCGTAGA GCTTGGGGGC
5461 GAGGAAGACG GACTCGGGGG CGTAGGCGTC CGCGCCGAG CTGGCGCAGA CCGTCTCGCA
5521 CTCCACGAGC CAGGTGAGGT CCGGGCGGTT GGGGTCAAAA ACGAGGTTTC CTCGTCGCTT
5581 TTTGATGCGT TTCTTACCTC TGGTCTCCAT GAGCTCGTGT CCGCGCTGGG TGACAAAGAG
5641 GCTGTCCGTG TCCCGTAGA CCGACTTTAT GGGCCGGTCC TCGAGCGGGG TGCCGCGGTC
5701 CTCGTCTAG AGGAACCCCG CCCACTCCGA GACGAAGGCC CGGGTCCAGG CCAGCACGAA
5761 GGAGGCCACG TGGGAGGGGT AGCGGTCTGT GTCCACCAGC GGGTCCACCT TCTCCAGGGT
5821 ATGCAAGCAC ATGTCCCCCT CGTCCACATC CAGGAAGGTG ATTGGCTTGT AAGTGTAGGC
5881 CACGTGACCG GGGGTCCCGG CCGGGGCGGT ATAAAAGGGG GCGGGCCCTT GCTCGTCTC
5941 ACTGTCTTCC GGATCGCTGT CCAGGAGCGC CAGCTGTTGG GGTAGGTATT CCTCTCGAA
6001 GCGGGGCATG ACCTCGGCAC TCAGGTTGTC AGTTTCTAGA AACGAGGAGG ATTTGATATT
6061 GACGGTGCCG TTGGAGACGC CTTTCATGAG CCCCTCGTCC ATTTGGTCAG AAAAGACGAT
6121 CTTTTTGTG TCGAGCTTGG TGGCGAAGGA GCCGTAGAGG GCGTTGGAGA GCAGCTTGGC
6181 GATGGAGCGC ATGGTCTGGT TCTTTTCCTT GTCGGCGCGC TCCTTGGCGG CGATGTTGAG
6241 CTGCACGTAC TCGCGCGCCA CGCACTTCCA TTCGGGGAAG ACGGTGGTGA GCTCGTCGGG
6301 CACGATTCTG ACCCGCCAGC CGCGGTTGTG CAGGGTGATG AGGTCCACGC TGGTGGCCAC
6361 CTCGCCGCGC AGGGGCTCGT TGGTCCAGCA GAGGCGCCCG CCTTGCAGG AGCAGAAGGG
6421 GGGCAGCGGG TCCAGCATGA GCTCGTCGGG GGGGTGCGG TCCACGGTGA AGATGCCGGG
6481 CAGGAGCTCG GGGTCCAAGT AGCTGATGCA GGTGCCAGA TTGTCCAGCG CCGCTTGCCA
6541 GTCGCGCACG GCCAGCGCGC GCTCGTAGGG GCTGAGGGG GTGCCCCAGG GCATGGGGTG
6601 CGTGAGCGCG GAGGCGTACA TGCCCGAGAT GTCGTAGACG TAGAGGGGCT CCTCGAGGAC
6661 GCCGATGTAG GTGGGGTAGC AGCGCCCCC CCGGATGCTG GCGCGCACGT AGTCGTACAG
6721 CTCGTGCGAG GCGCGAGGA GCGCCGTGCC GAGGTTGGAG CGTTGCGGCT TTTGCGGCGC
6781 GTAGACGATC TGGCGGAAGA TGGCGTGGGA GTTGGAGGAG ATGGTGGGCC TTTGGAAGAT
6841 GTTGAAGTGG GCGTGGGGCA GGCCGACCGA GTCCCTGATG AAGTGGGCGT AGGAGTCCTG
6901 CAGCTTGGCG ACGAGCTCGG CCGTGACGAG GACGTCCAGG GCGCAGTAGT CGAGGGTCTC
6961 TTGGATGATG TCATACTTGA GCTGGCCCTT CTGCTTCCAC AGCTCGCGGT TGAGAAGGAA
7021 CTCTTCGCGG TCCTTCCAGT ACTCTTCGAG GGGGAACCCG TCCTGATCGG CACGGTAAGA
7081 GCCCACCATG TAGAAGTGGT TGACGCGCTT GTAGGCGCAG CAGCCCTTCT CCACGGGGAG
7141 GCGGTAAGCT TGCGCGGCCT TGCGCAGGGA GGTGTGGGTG AGGGCGAAGG TGTCGCGCAC
7201 CATGACCTTG AGGAAGTGGT GCTTGAAGTC GAGGTGCTCG CAGCCGCCCT GCTCCAGAG

Fig. 7B

7261 TTGGAAGTCC GTGCGCTTCT TGTAGGCGGG GTTAGGCAAA GCGAAAGTAA CATCGTTGAA
7321 GAGGATCTTG CCCGCGCGGG GCATGAAGTT GCGAGTGATG CCGAAAGCGT GGGGCACCTC
7381 GGCCCGGTTG TTGATGACCT GGGCGGCGAG GACGATCTCG TCGAAGCCGT TGATGTTGTG
7441 CCCGACGATG TAGAGTTCCA CGAATCGCGG GCGGCCCTTG ACGTGGGGCA GCTTCTTGAG
7501 CTCGTCGTAG GTGAGCTCGG CGGGGTGCGT GAGCCCGTGC TGCTCGAGGG CCCAGTCGGC
7561 GACGTGGGGG TTGGCGCTGA GGAAGGAAGT CCAGAGATCC ACGGCCAGGG CGGTCTGCAA
7621 GCGGTCCCGG TACTGACGGA ACTGTTGGCC CACGGCCATT TTTTCGGGGG TGACGCAGTA
7681 GAAGGTGCGG GGGTCGCCGT GCCANCGGT CCACTTGAGC TGGAGGGCGA GGTCGTGGGC
7741 GAGCTCGACG AGCGCGGGT CCCCAGAGAG TTTCATGACC AGCATGAAGG GGACGAGCTG
7801 CTTGCCGAAG GACCCCATCC AGGTGTAGGT TTCCACATCG TAGGTGAGGA AGAGCCTTTC
7861 GGTGCGAGGA TGCGAGCCGA TGGGAAGAA CTGGATCTCC TGCCACCAGT TGGAGGAATG
7921 GCTGTTGATG TGATGGAAGT AGAAATGCCG ACGGCGCGCC GAGCACTCGT GCTTGTGTTT
7981 ATACAAGCGT CCGCAGTGCT CGCAACGCTG CACGGGATGC ACGTGCTGCA CGAGCTGTAC
8041 CTGGGTTCCCT TTGGCGAGGA ATTTCACTGG GCAGTGGAGC GCTGGCGGCT GCATCTCGTG
8101 CTGTACTACG TCTTGCCCAT CGGCGTGGCC ATCGTCTGCC TCGATGGTGG TCATGCTGAC
8161 GAGCCCGCGC GGGAGGACG TCCAGACCTC GGCTCGGACG GGTGCGAGAG CGAGGACGAG
8221 GCGCGCAGG CCGGAGCTGT CCAGGTCCTT GAGACGCTGC GGAGTCAGGT CAGTGGGCAG
8281 CCGCGGCGCG CGGTTGACTT GCAGGAGCTT TTCCAGGGCG CGCGGGAGGT CCAGATGGTA
8341 CTGATCTCC ACGGCGCCGT TGGTGGCTAC GTCCACGGCT TGCAGGGTGC CGTGCCCTG
8401 GGGCGCCACC ACCGTGCCCC GTTCTTCTT GGGCGCTGCT TCCATGTCGG TCAGAAGCGG
8461 CCGCGAGGAC GCGCGCCGGG CCGCAGGGSC GGCTCGGGGC CCGGAGGCAG GGGCGCAGG
8521 GGCACGTCGG CGCCGCGCGC GGGCAGGTTT TGGTACTGCG CCCGAGAAG ACTGCGCTGA
8581 GCGACGACGC GACGGTTGAC GTCTTGATC TGACGCCTCT GGGTGAAGGC CACGGGACCC
8641 GTGAGTTTGA ACCTGAAAGA GAGTTCGACA GAATCAATCT CCGTATCGTT GACGGCGGCC
8701 TGCCGAGGA TCTCTTGAC GTCGCCGAG TTGTCTGGT AGGCGATCTT GGTATGAAC
8761 TGCTCGATCT CCTCCTCTG AAGGTCTCCG CCGCCGGCGC GCTCGACGGT GGCCGCGAGG
8821 TCGTTGGAGA TGCGGCCCAT GAGCTCGAG AAGGCGTTCA TGCCGGCCTC GTTCCAGACG
8881 CCGCTGTAGA CCACGGCTCC GTCCGGGTGCG CGCGCGCGCA TGACCACCTG GCGAGGTTG
8941 AGCTCGACGT GCGCGGTGAA GACCGCGTAG TTGCAGAGGC GCTGGTAGAG GTAGTTGAGC
9001 GTGGTGGCGA TGTGCTCGGT GACGAAGAAG TACATGATCC AGCGGCGGAG CGGCATCTCG
9061 CTGACGTCGC CCAGGGCTTC CAAGCGTTC ATGGCTCTGT AGAAGTCCAC GCGAAGTTG
9121 AAAAAGTGG AGTTGCGCGC CGAGACGGTC AACTCCTCCT CCAGAAGACG GTCGCAAGGT
9181 GCGATGGTGG CGCGCACCTC GCGCTCGAAG GCGCGGGGG GCTCCTCTTC CATCTCCTCC
9241 TCTTCTCCT CCACTAACAT CTCTTCTACT TCCTCCTCAG GAGGCGGTGG CCGGGGAGGG
9301 GCCCTGCGTC GCCGCGGCG CACGGGCAGA CCGTGCATGA AGCGCTCGAT GGTCTCCCCG
9361 CGCCGGCGAC GCATGGTCTC GGTGACGGCG CGCCCGTCTT CGCGGGGCGG CAGCATGAAG
9421 ACGCCGCGCG GCATCTCCAG GTGGCCGCGG GGGGGGTCTC CGTTGGGCAG GGAGAGGGCG
9481 CTGACGATGC ATCTTATCAA TTGACCCGTA GGGACTCCGC GCAAGGACCT GAGCGTCTCG
9541 AGATCCACGG GATCCGAAAA CCGCTGAACG AAGGCTTCGA GCCAGTCGCA GTCGCAAGGT
9601 AGGCTGAGCC CGGTTTCTTG TTCTTCGGGT ATTTGGTTCGG GAGGCGGCGG GCGATGCTGC
9661 TGGTGATGAA GTTGAAGTAG GCGGTCTGTA GACGGCGGAT GGTGGCGAGG AGCACCAGGT
9721 CTTGGGCCCC GGCTTGCTGG ATGCGCAGAC GGTCGGCCAT GCGCCAGGCG TGGTCTTGAC
9781 ACCTGGCGAG GTCTTGTAG TAGTCTGCA TGAGCCGCTC CACGGGCACC TCCTCCTCGC
9841 CCGCGCGGCC GTGCATGCGC GTGAGCCCGA ACCGCGCTG CCGCTGGACG AGCGCCAGGT
9901 CCGCGACGAC GCGCTCGGTG AGGATGGCCT GCTGGATCTG GGTGAGGGTG GTCTGGAAGT
9961 CGTCGAAGTC GACGAAGCGG TGGTAGGCTC CGGTGTTGAT GGTGTAGGAG CAGTTGGCCA
10021 TGACGACCA GTTGACGGTC TGGTGGCCGG GTCGCACGAG CTCGTGGTAC TTGAGGCGCG
10081 AGTAGGCGCG CGTGTCGAAG ATGTAGTCGT TGCAGGCGCG CACGAGGTAC TGGTATCCGA
10141 CGAGGAAGTG CCGCGGCGGC TGGCGGTAGA GCGGCCATCG CTCGGTGGCG GGGGCGCCGG
10201 GCGCGAGGTC CTCGAGCATG AGGCGGTGGT AGCCGTAGAT GTACCTGGAC ATCCAGGTGA
10261 TGCCGGCGCG GGTGGTGGAG GCGCGCGGGA ACTCGCGGAC GCGGTTCCAG ATGTTGCGCA
10321 GCGCGAGGAA GTAGTTCATG GTGGCCGCGG TCTGGCCCGT GAGGCGCGCG CAGTCGTGGA
10381 TGCTCTAGAC ATACGGGCAA AAACGAAAGC GGTACGCGC TCGACTCCGT GGCCTGGAGG
10441 CTAAGCGAAC GGGTTGGGCT GCGCGTGTAC CCCGGTTCGA ATCTCGAATC AGGCTGGAGC
10501 CGCAGCTAAC GTGGTACTGG CACTCCCGTC TCGACCCAAG CCTGCTAACG AAACCTCCAG
10561 GATACGGAGG CGGGTCTGTT TTTGGCCTTG GTCGCTGGTC ATGAAAACT AGTAAGCGCG
10621 GAAAGCGGCC GCGCGCATG GCTCGCTGCC GTAGTCTGGA GAAAGAATCG CCAGGTTGTC
10681 GTTGGCGGTG GCGCGGTTT GAGCCTCAGC GCTCGGCGCC GGCCGGATTG CGCGGCTAAC
10741 GTGGGCGTGG CTGCCCCGTC GTTCCCAAGA CCCCTTAGCC AGCCGACTTC TCCAGTTACG
10801 GAGCGAGCCC CTCTTTTTTT TTCTTGTGTT TTTGCCAGAT GCATCCCGTA CTGCGGCAGA

Fig. 7C

10861 TGCGCCCCCA CCCTCCACCA CAACCGCCCC TACCGCAGCA GCAGCAACAG CCGGCGCTTC
10921 TGCCCCCGCC CCAGCAGCAG CCAGCCACTA CCGCGGCGGC CGCCGTGAGC GGAGCCGGCG
10981 TTCAGTATGA CCTGGCCTTG GAAGAGGGCG AGGGGCTGGC GCGGCTGGGG CCGTCTGCGC
11041 CGGAGCGGCA CCCGCGCGTG CAGATGAAAA GGGACGCTCG CGAGGCCTAC GTGCCCAAGC
11101 AGAACCTGTT CAGAGACAGG AGCGGCGAGG AGCCCCAGGA GATGCGCGCC TCCCCGCTTC
11161 ACGCGGGGCG GGAGCTGCGG CGCGGCCTGA ACCGAAAGCG GGTGCTGAGG GACGAGGATT
11221 TCGAGGCGGA CGAGCTGACG GGGATCAGCC CCGTGCGCGC GCACGTGGTC GNGNCAACC
11281 TGGTCACGGC GTACGAGCAG ACCGTGAAGG AGGAGAGCAA CTTCACAAAA TCCTTCAACA
11341 ACCACGTGCG ACCTTGATC GCGCGCGAGG AGGTGACCTT GGGCCTGATG CACCTGTGGG
11401 ACCTGCTGGA GGCCATCGTG CAGAACCCCA CGAGCAAGCC GCTGACGGCG CAGCTGTTTC
11461 TGGTGGTGCA GCACAGTCGG GACAACGAGA CGTTCAGGGA GGCCTGCTG AATATCACCG
11521 AGCCCCAGGG CCGCTGGCTC CTGGACCTGG TGAACATTTT GCAGAGCATC GTGGTGACAG
11581 AGCGCGGGCT GCCGCTGTCC GAGAAGCTGG CGGCCATCAA CTTCGCGGTG CTGAGTCTGG
11641 GCAAGTACTA CGCTAGGAAG ATCTACAAGA CCCCCTACGT GCCCATAGAC AAGGAGGTGA
11701 AGATCGACGG GTTTTACATG CGCATGACCC TGAAAGTGCT GACCCTGAGC GACGATCTGG
11761 GGGTGATACG CAACGACAGG ATGCACCGCG CGGTGAGCGC CAGCCGCCGG CGCGAGCTGA
11821 GCGACCAGGA GCTGATGCAC AGCCTGCAGC GGGCCCTGAC CGGGGCCGGG ACCGAGGGGG
11881 AGAGCTACTT TGACATGGGC GCGGACCTGC GCTGGCAGCC CAGCCGCCGG GCCTTGGAAG
11941 CTGCCGCGCG TTCCCCCTAC GTGGAGGAGG TGGACGATGA GGAGGAGGAG GGCGAGTACC
12001 TGGAAGACTG ATGGCGCGAC CGTATTTTTG CTAGATGCAG CAACAGCCAC CGCCGCGGCC
12061 TCCTGATCCC GCGATGCGGG CGGCGCTGCA GAGCCAGCCG TCCGGCATT AACTCTCGGA
12121 CGATTGGACC CAGGCCATGC AACGCATCAT GCGCTGACG ACCCGCAATC CCGAAGCTTT
12181 TAGACAGCAG CCTCAGGCCA ACCGGCTCTC GGCCATCCTG GAGGCCGTGG TGCCCTCGCG
12241 CTCGAACCCC ACGCACGAGA AGGTGCTGGC CATCGTGAAC GCGCTGGTGG AGAACAAGGC
12301 CATCCGCGGT GACGAGGCCG GGCTGGTGTA CAACGCGCTG CTGGAGCGCG TGGCCCCTTA
12361 CAACAGCACC AACGTGCAGA CGAACCTGGA CCGCATGGTG ACCGACGTGC GCGAGGCGGT
12421 GTCGCAGCGC GAGCGGTTC ACCGCGAGTC GAACCTGGGC TCCATGGTGG CGCTGAACGC
12481 CTTCTGAGC ACGCAGCCG CCAACGTGCC CCGGGGCCAG GAGGACTACA CCAACTTCAT
12541 CAGCGCGCTG CCGCTGATGG TGGCCGAGGT GCCCAGAGC GAGGTGTACC AGTCGGGCC
12601 GGACTACTTC TTCCAGACCA GTCGCCAGGG CTTGCAGACC GTGAACCTGA GCCAGGCTTT
12661 CAAGAACTTG CAGGGACTGT GGGGCGTGCA GGCCCCGGTC GGGGACCGCG CGACGGTGTC
12721 GAGCCTGCTG ACGCCGAAC TCGCCTGTCT GCTGCTGCTG GTGGCGCCCT TCACGGACAG
12781 CGGCAGCGTG AGCCGCGACT CGTACCTGGG CTACCTGCTT AACCTGTACC GCGAGGCCAT
12841 CGGACAGGCG CACGTGGACG AGCAGACCTA CCAGGAGATC ACCCAGTGA GCCGCGCGCT
12901 GGGCCAGGAG GACCCGGGCA ACCTGGAGGC CACCTGAAC TTCTTGCTGA CCAACCGGTC
12961 GCAGAAGATC CCGCCCCAGT ACGCGCTGAG CACCGAGGAG GAGCGCATCC TCGCTACGT
13021 GCAGCAGAGC GTGGGGCTGT TCCTGATGCA GGAGGGGGCC ACGCCAGCG CGGCGCTCGA
13081 CATGACCGCG CGCAACATGG AGCCCAGCAT GTACGCCCGC AACCGCCGT TCATCAATAA
13141 GCTGATGGAC TACTTGATC GGGCGGCCGC CATGAACTCG GACTACTTTA CCAACGCCAT
13201 CTTGAACCCG CACTGGCTCC CGCCGCCCCG GTTCTACACG GCGGAGTACG ACATGCCCGA
13261 CCCCACGAC GGGTTCTGT GGGACGACGT GGACAGCAGC GTGTTCTCGC CGCGTCCAGG
13321 AACCAATGCC GTGTGGAAGA AAGAGGGCGG GGACCGCGCG CCGTCTCGG CGCTGTCCGG
13381 TCGCGCGGGT GCTGCCGCGG CGGTGCCCGA GGCCGCCAGC CCCTTCCCGA GCCTGCCCTT
13441 TTCGCTGAAC AGCGTGCGCA GCAGCGAGCT GGTGCGGCTG ACGCGACCGC GCCTGTGCGG
13501 CGAGGAGGAG TACCTGAACG ACTCCTTGTT GAGGCCCGAG CGCGAGAAGA ACTTCCCCAA
13561 TAACGGGATA GAGAGCCTGG TGGACAAGAT GAGCCGCTGG AAGACGTACG CGCACGAGCA
13621 CAGGGACGAG CCCCAGCTA GCAGCGCAGG CACCCGTAGA CGCCAGCGGC ACGACAGGCA
13681 GCGGGGACTG GTGTGGGACG ATGAGGATTC CGCCGACGAC AGCAGCGTGT TGGACTTGGG
13741 TGGGAGTGGT GGTAACCCGT TCGCTCACCT GCGCCCCGT ATCGGGCGCC TGATGTAAGA
13801 ATCTGAAAAA ATAAAAGACG GTACTACCA AGGCCATGGC GACCAGCGTG CGTTCCTCTC
13861 TGTTGTTTGT AGTAGTATGA TGAGGCGCGT GTACCCGGAG GGTCTCTCTC CCTCGTACGA
13921 GAGCGTGATG CAGCAGGCGG TGGCGGCGG GATGCAGCCC CCGCTGGAGG CGCCTTACGT
13981 GCCCCGCGG TACCTGGCGC CTACGGAGGG GCGGAACAGC ATTCTGTACT CCGAGCTGGC
14041 ACCCTTGTAC GATACCACCC GGTGTGACCT GGTGGACAAC AAGTCGGCAG ACATCGCCTC
14101 GCTGAACCTAC CAGAACGACC ACAGCAACTT CCTGACCACC GTGGTGACAGA ACAACGATTT
14161 CACCCCCACG GAGGCCAGCA CCCAGACCAT CAACCTTGAC GAGCGCTCGC GGTGGGGCGG
14221 CCAGCTGAAA ACCATCATGC ACACCAACAT GCCCAACGTG AACGAGTCA TGTACAGCAA
14281 CAAGTTCAAG GCGCGGGTGA TGGTCTCGCG CAAGACCCCC AACGGGGTGG ATGATGATTA
14341 TGATGGTAGT CAGGACGAGC TGACCTACGA GTGGGTGGAG TTTGAGCTGC CCGAGGGCAA
14401 CTTCTCGGTG ACCATGACCA TCGATCTGAT GAACAACGCC ATCATCGACA ACTACTTGGC

Fig. 7D

14461 GGTGGGGCGG CAGAACGGGG TGCTGGAGAG CGACATCGGC GTGAAGTTCG ACACGCGCAA
14521 CTTCCGGCTG GGCTGGGACC CCGTGACCGA GCTGGTGATG CCGGGCGTGT ACACCAACGA
14581 GGCCCTCCAC CCCGACATCG TCCTGTGTCG CGGCTGCGGC GTGGACTTCA CCGAGAGCCG
14641 CCTCAGCAAC CTGCTGGGCA TCCGCAAGCG GCAGCCCTTC CAGGAGGGCT TCCAGATCCT
14701 GTACGAGGAC CTGGAGGGGG GCAACATCCC CGCGCTCTTG GATGTGCAAG CCTACGAGAA
14761 AAGCAAGGAG GATAGCACCG CCGCGGCGAC CGCAGCCGTG GCCACCGCCT CTACCGAGGT
14821 GCGGGGCGAT AATTTTGCTA GCGCTGCGGC AGCGGCCGAG GCGGCTGAAA CCGAAAAGTAA
14881 GATAGTCATC CAGCCGGTGG AGAAGGACAG CAAGGACAGG AGCTACAACG TGCTCGCGGA
14941 CAAGAAAAAC ACCGCCTACC GCAGCTGGTA CCTGGCCTAC AACTACGGCG ACCCCGAGAA
15001 GGGCGTGCGC TCCTGGACGC TGCTCACCAC CTCGGACGTC ACCTGCGGCG TGGAGCAAGT
15061 CTACTGGTCG CTGCCCGACA TGATGCAAGA CCCGGTCACC TTCCGCTCCA CGCGTCAAGT
15121 TAGCAACTAC CCGGTGGTGG GCGCCGAGCT CCTGCCCGTC TACTCCAAGA GCTTCTTCAA
15181 CGAGCAGGCC GTCTACTCGC AGNAGCTGCG CGCCTTCACC TCGCTCACGC ACGTCTTCAA
15241 CCGCTTCCCC GAGAACCAGA TCCTCGTCCG CCGCCGCGCC CACCATFACC ACCGTCAAGT
15301 AAAACGTTCC TGCTCTCACA GATCAGGGA CCTGCCGCT CCGCAGCAGT ATCCGGGGAG
15361 TCCAGCGCGT GACCGTCACT GACGCCAGAC GCCGCACCTG CCCCTACGTC TACAAGGCCC
15421 TGGGCGTAGT CGCGCCGCGC GTCCTCTCGA GCCGCACCTT CTAAAAAATG TCCATTCTCA
15481 TCTCGCCAG TAATAACACC GGTGGGGGCC TGCGCGCGCC CAGCAAGATG TACGAGGGCG
15541 CTCGCCAACG CTCCACGCAA CACCCCGTGC GCCTGCGCGG GCACTTCCGC GCTCCCTGGG
15601 GCGCCCTCAA GGGCCGCGTG CGCTCGCGCA CCACCGTCGA CGACGTGATC GACCAGGTGG
15661 TGGCCGACGC GCGCAACTAC ACGCCCGCCG CCGCGCCCGT CTCCACCGTG GACGCCGTCA
15721 TCGACAGCGT GGTGGCCGAC GCGCGCCGGT ACGCCCGCAC CAAGAGCCGG CGGCGGCGCA
15781 TCGCCCGCGG GCACCGGAGC ACCCCCGCCA TGCGCGCGGC GCGAGCCTTG CTGCGCAGGG
15841 CCAGGCGCAC GGGACGCGAG GCCATGCTCA GGGCGGCCAG ACGCGCGGCC TCCGGCAGCA
15901 GCAGCGCCGG CAGGACCCGC AGACGCGCGG CCACGGCGGC GGGCGCGGCC ATCGCCAGCA
15961 TGTCCCGCCC GCGGCGCGGC AACGTGTACT GGGTGCGCGA CGCCGCCACC GGTGTGCGCG
16021 TGCCCGTGCG CACCCGCCCC CCTCGCACTT GAAGATGCTG ACTTCGCGAT GTTGATGTGT
16081 CCCAGCGGCG AGGAGGATGT CCAAGCGCAA ATACAAGGAA GAGATGCTCC AGGTCATCGC
16141 GCCTGAGATC TACGGCCCCG CGGCGCGGGT GAAGGAGGAA AGAAAGCCCC GCAAACGTAA
16201 GCGGGTCAAA AAGGACAAAA AGGAGGAGGA AGATGACGGA CTGGTGGAGT TTGTGCGCGA
16261 GTTCGCCCCC CGGCGGCGCG TGCACTGGCG CGGGCGGAAA GTGAAACCGG TGCTGCGGCC
16321 CGGCACCACG GTGGTCTTCA CGCCCGCGCA GCGTTCCGGC TCCGCCTCCA AGCGCTCCTA
16381 CGACGAGGTG TACGGGGACG AGGACATCCT CGAGCAGGCG GTCGAGCGTC TGGGCGAGTT
16441 TGCGTACGGC AAGCGCAGCC GCCCGCGGCC CTTGAAAGAG GAGGCGGTGT CCATCCCGCT
16501 GGACCACGGC AACCACACGC CGAGCGTGAA GCCGGTGACC CTGCAGCAGG TGCTACCGAG
16561 CGCGGCGCGG CGCCGGGGCT TCAAGCGCGA GGGCGGCGAG GATCTGTACC CGACCATGCA
16621 GCTGATGGTG CCCAAGCGCC AGAAGCTGGA GGACGTGCTG GAGCACATGA AGGTGAGACC
16681 CGAGGTGCAG CCCGAGGTCA AGGTGCGGCC CATCAAGCAG GTGGCCCCGG GCCTGGGCGT
16741 GCAGACCGTG GACATCAAGA TCCCCACGGA GCCCATGGAA ACGCAGACCG AGCCCGTGAA
16801 GCCAGCACCC AGCACCATGG AGGTGCAGAC GGATCCCTGG ATGCCAGCAC CAGCTTCCAC
16861 CAGCACTCGC CGAAGACGCA AGTACGGCGC GGCCAGCCTG CTGATGCCCA ACTACGCGGC
16921 TGCATCCTTC CATCATCCCC ACGCCGGGCT ACCGCGGCAC GCGCTTCTAC CGCGGCTACA
16981 CCAGCAGCCG CCGCCGCAAG ACCACCACCC GCGCCGTCG TCGCAGCCGC CGCAGCAGCA
17041 CCGCGACTTC CGCCTTGGTG CGGAGAGTGT ATCGCAGCGG GCGCGAGCCT CTGACCTGCG
17101 CGCGCGCGCG CTACCACCCG AGCATCGCCA TTAACTACC GCCTCCTACT TGCAGATATG
17161 GCCCTCACAT GCCGCTCCG CGTCCCCATT ACGGGCTACC GAGGAAGAAA GCCGCGCCGT
17221 AGAAGGCTGA CGGGGAACGG GCTGCGTCGC CATCACCACC GCGGCGGCG CGCCATCAGC
17281 AAGCGGTTGG GGGGAGGCTT CCTGCCCGCG CTGATCCCCA TCATCGCCGC GCGATCGGG
17341 GCGATCCCCG GCATAGCTTC CGTGGCGGTG CAGGCCTCTC AGCGCCACTG AGACACAAAA
17401 AAGCATGGAT TTGTAATAAA AAAAAAATG GACTGACGCT CCTGGTCTTG TGATGTGTGT
17461 TTTTAGATGG AAGACATCAA TTTTTCGTCC CTGGCACCGC GACACGGCAC CCGGCCGTTT
17521 ATGGGCACCT GGAGCGACAT CGGCAACAGC CAACTGAACG GGGGCGCCTT CAATTGGAGC
17581 AGTCTCTGGA GCGGGCTTAA GAATTTGCGG TCCACGCTCA AAACCTATGG CAACAAGGCG
17641 TGGAACAGCA GCACAGGGCA GCGCTGAGG GAAAAGCTGA AAGAACAGAA CTTCCAGCAG
17701 AAGGTGGTTG ATGGCCTGGC CTCAGGCATC AACGGGGTGG TTGACCTGGC CAACCAGGCC
17761 GTGCAGAAAC AGATCAACAG CCGCTGGAG CCGGTCCCGC CCGCGGGGTC CGTGGAGATG
17821 CCCCAGGTGG AGGAGGAGCT GCCTCCCTG GACAAGCGCG GCGACAAGCG ACCGCGTCCC
17881 GACGCGGAGG AGACGCTGCT GACGCACAG GACGAGCCGC CCCCCTACGA GGAGGCGGTG
17941 AAAC TGGGCC TGCCACAC GCGGCCCCG GCGCCTCTGG CCACCGGAGT GCTGAAACCC
18001 AGCAGCAGCC AGCCCCGCGAC CCTGGACTTG CCTCCGCTC GCCCTCCAC AGTGGCTAAG

Fig. 7E

18061 CCCCTGCCGC CGGTGGCCGT CGCGTCGCGC GCGCCCCGAG GCGCCCCCA GGCGAACTGG
18121 CAGAGCACTC TGAACAGCAT CGTGGGTCTG GGAGTGCAGA GTGTGAAGCG CCGCCGCTGC
18181 TATTAAGA CACTGTAGCG CTTAACTTGC TTGTCTGTGT GTATATGTAT GTCCGCCGAC
18241 CAGAAGGAGG AGTGTGAAGA GGCGCGTCGC CGAGTTGCAA GATGGCCACC CCATCGATGC
18301 TGCCCCAGTG GCGGTACATG CACATCGCCG GACAGGACGC TTCGGAGTAC CTGAGTCCGG
18361 GTCTGGTGCA GTTCGCCCCG GCCACAGACA CCTACTTCAG TCTGGGGAAC AAGTTTAGGA
18421 ACCCCACGGT GCGGCCACG CACAATGTGA CCACCGACCG CAGCCAGCGG CTGACGGTGC
18481 GCTTCGTGCC CGTGGACCGC GAGGACAACA CCTACTCGTA CAAAGTGCGC TACACGCTGG
18541 CCGTGGGCGA CAACCGCGTG CTGGACATGG CCAGCACCTA CTTTGACATC CGCGGCGTGC
18601 TGGACCGGG CCCTAGCTTC AAACCTACT CTGGCACCGC CTACAACAGC CTAGCTCCCA
18661 AGGGAGCTCC CAATTCCAGC CAGTGGGAGC AAGCAAAAC AGGCAATGGG GGAATATGG
18721 AAACACACAC ATATGGTGTG GCCCCAATGG GCGGAGAGAA TATTACAAA GATGGTCTTC
18781 AAATTGGAAC TGACGTTACA GCGAATCAGA ATAAACCAAT TTATGCCGAC AAAACATTTT
18841 AACCAGAACC GCAAGTAGGA GAAGAAATT GGCAAGAAAC TGAAACTTT TATGGCGGTA
18901 GAGCTCTTAA AAAAGACACA AACATGAAAC CTTGCTATGG CTCCTATGCT AGACCCACCA
18961 ATGAAAAGG AAGTCAAGCT AAACCTAAG TTGGAGATGA TGGAGTTCCA ACCAAAGAAT
19021 TCGACATAGA CCTGGCTTTC TTTGATATC CCGGTGGCAC CGTGAACGGT CAGACGAGT
19081 ATAAAGCAGA CATTGTCATG TATACCGAAA ACACGTATTT GGAAACTCCA GACACGCATG
19141 TGGTATACAA ACCAGGCAAG GATGATGCAA GTTCTGAAAT TAACCTGGTT CAGCAGTCTA
19201 TGCCCAACAG ACCCAACTAC ATTGGGTTC GGGACAACCT TATCGGTCTT ATGTACTACA
19261 ACAGCACTGG CAATATGGGT GTGCTTGCTG GTCAGGCCTC CCAGCTGAAT GCTGTGGTTG
19321 ATTTGCAAGA CAGAAACACC GAGCTGTCTT ACCAGCTCTT GCTTGACTCT TTGGGTGACA
19381 GAACCCGGTA TTTCAGTATG TGAACCCAG CCGTGGACAG TTATGACCCC GATGTGCGCA
19441 TCATCGAAAA CCATGGTGTG GAGGATGAAT TGCCAAACTA TTGCTTCCCC TTGGACGGCT
19501 CTGGCACTAA CGCCGCATAC CAAGGTGTGA AAGTAAAAGA TGGTCAAGAT GGTGATGTTG
19561 AGAGTGAATG GGAAATGAC GATACTGTTG CAGCTCGAAA TCAATTATGT AAAGGTAACA
19621 TTTTCGCCAT GGAGATTAAT CTCCAGGCTA ACCTGTGGAG AAGTTTCCTC TACTCGAACG
19681 TGGCCCTGTA CCTGCCCGAC TCCTACAAGT ACACGCCGAC CAACGTCACG CTGCCGACCA
19741 ACACCAACAC CTACGATTAC ATGAATGGCA GAGTGACACC TCCCTCGCTG GTAGACGCCT
19801 ACCTCAACAT CGGGGCGCGC TGGTCGCTGG ACCCATGGA CAACGTCAAC CCCTTCAACC
19861 ACCACCGCAA CGCGGCGCTG CGCTACCGCT CCATGCTCCT GGCACAACGG GGTGACGGT
19921 CCTTCCACAT CCAGGTGCCC CAAAAGTTTT TCGCCATCAA GAGCCTCCTG CTCCTGCCCG
19981 GGTCTTACAC CTACGAGTGG AACTTCCGCA AGGACGTCAA CATGATCCTG CAGAGCTCCC
20041 TAGGCAACGA CCTGCGCACG GACGGGGCCT CCATCGCCTT CACCAGCATC AACCTCTACG
20101 CCACCTTCTT CCCCATGGCG CACAACACCG CCTCCACGCT CGAGGCCATG CTGCGCAACG
20161 ACACCAACGA CAGTCTCTC AACGACTACC TCTCGGCGGC CAACATGCTC TACCCCATCC
20221 CGGCCAACGC CACCAACGTG CCCATCTCCA TCCCCTCGCG CAACTGGGCG GCCTTCCGCG
20281 GATGGTCTCT CACGCGCCTG AAGACCCGCG AGACGCCCTC GCTCGGCTCC GGTTCGACC
20341 CCTACTTCGT CTACTCGGGC TCCATCCCCT ACCTAGACGG CACCTTCTAC CTCAACCACA
20401 CCTTCAAGAA GGTCTCCATC ACCTTCGACT CCTCCGTCAG CTGGCCCGGC AACGACCGCC
20461 TCCTGACGCC CAACGAGTTC GAAATCAAGC GCACCGTCGA CGGAGAGGGA TACAACGTGG
20521 CCCAGTGCAA CATGACCAAG GACTGGTTCC TGGTCCAGAT GCTGGCCAC TACAACATCG
20581 GCTACAGGG CTCTACGTG CCCGAGGGT ACAAGGACCG CATGTACTCC TTCTTCCGCA
20641 ACTTCCAGCC CATGAGCCG CAGTCTGTG ACGAGGTCAA CTACAAGGAC TACCAAGCCG
20701 TCACCTTGGC CTACCAGCAC AACAACCTCG GCTTCGTCGG CTACCTCGCG CCCACCATG
20761 GCCAGGGCCA GCCCTACCCC GCCAACCTACC CCTACCCGCT CATCGGCAAG AGCGCCGTCG
20821 CCAGCGTCAC CCAGAAAAAG TTCCTCTGCG ACCGGGTCAT GTGGCGCATC CCCTTCTCCA
20881 GCAACTTCAT GTCCATGGGC GCGCTCACCG ACCTCGGCCA GAACATGCTC TACGCCAACT
20941 CCGCCACGCG GCTAGACATG AATTTGCAAG TCGACCCCAT GGATGAGTCC ACCCTTCTCT
21001 ATGTTGTCTT CGAAGTCTTC GACGTCTGTC GAGTGCACCA GCGCCACCGC GCGCTCATCG
21061 AAGCCGTCTA CTGCGCACG CCTTCTCTCG CCGCAACGC CACCACCTAA GCGCTCTTTC
21121 CTCTTTGCAA GATGACGGCG GGCTCCGGCG ACGAGGAGCT CAGGCGCATC CTCCGCGACC
21181 TGGGCTGCGG GCCCTGCTTC CTGGGCACCT TCGACAAGCG CTTCCCTGGA TTCATGGCCC
21241 CGCACAAGCT GGCCTGCGCC ATCGTGAACA CGGCCGGCCG CGAGACCGGG GCGGAGCACT
21301 GGCTGGCCTT CGCCTGGAAC CCGCGCTCCC ACACATGCTA CCTCTTCGAC CCCTTCGGGT
21361 TCTCGGACGA GCGCCTCAAG CAGATCTACC AGTTCGAGTA CGAGGGCCTG CTGCGTCGCA
21421 GCGCCCTGGC CACCGAGGAC CGCTGCGTCA CCCTGGAAAA GTCCACCCAG ACCGTGCAGG
21481 GTCCGCGCTC GCGCGCTGCT GGGCTCTTCT GCTGCATGTT CCTGCACGCC TTCGTGCACT
21541 GGCCCGACCG CCCCATGGAC AAGAATCCCA CCATGAACCT ACTGACGGGG GTGCCCAACG
21601 GCATGCTCCA GTCGCCCCAG GTGGAACCCA CCCTGCGCCG CAACAGGAA GCGCTCTACC

Fig. 7F

21661 GCTTCCTCAA TGCCCACTCC GCCTACTTTC GCTCCCACCG CGCGCGCATC GAGAAGGCCA
21721 CCGCCTTCGA CCGCATGAAT CAAGACATGT AAAAAACCGG TGTGTGTATG TGAATGCTTT
21781 ATTCATAATA AACAGCACAT GTTTATGCCA CCTTCTCTGA GGCTCTGACT TTATTTAGAA
21841 ATCGAAGGGG TTCTGCCGGC TCTCGGCATG GCCCGCGGGC AGGGATACGT TGCGGAACTG
21901 GTACTTGGGC AGCCACTTGA ACTCGGGGAT CAGCAGCTTG GGCACGGGGA GGTCGGGGAA
21961 CGAGTCGCTC CACAGCTTGC GCGTGAGTTG CAGGGCGCCC AGCAGGTCCG GCGCGGAGAT
22021 CTTGAAATCG CAGTTGGGAC CCGCGTCTTG CGCGCGAGAG TTGCGGTACA CGGGGTTGCA
22081 GCACTGGAAC ACCATCAGGG CCGGGTGCTT CACGCTTGCC AGCACCGTCG CGTCGGTGAT
22141 GCCCTCCACG TCCAGATCCT CGGCGTTGGC CATCCCGAAG GGGGTCATCT TGCAGGTCTG
22201 CCGCCCCATG CTGGGCACGC AGCCGGGCTT GTGGTTGCAA TCGCAGTGCA GGGGGATCAG
22261 CATCATCTGG GCCTGCTCGG AGCTCATGCC CCGGTACATG GCCTTCATGA AAGCCTTCAG
22321 CTGGCGGAAG GCCTGCTGCG CTTGCGCCG CTCGGTGAAG AAGACCCCGC AGGACTTGCT
22381 AGAGAACTGG TTGGTGGCGC AGCCGGCGTC GTGCACGCAG CAGCGCGCGT CGTTGTTGGC
22441 CAGCTGCACC ACGCTGCGCC CCCAGCGTCT CTGGGTGATC TTGGCCCGGT TGGGGTTCTC
22501 CTTCAGCGCG CGCTGCCCGT TCTCGCTCGC CACATCCATC TCGATAGTGT GCTCCTTCTG
22561 GATCATCACG GTCCCGTGCA GGCACCGCAG CTTGCCCTCG GCTTCGTTGC AGCCGTGCAG
22621 CCACAGCGCG CAGCCGGTGC ACTCCAGTT CTTGTGGGCG ATCTGGGAGT GCGAGTGCAC
22681 GAAGCCCTGC AGGAAGCGGC CCATCATCGC GGTCAGGGTC TTGTTGCTGG TGAAGGTCAG
22741 CGGGATGCCG CGGTGCTCCT CGTTCACATA CAGGTGGCAG ATGCGCGCGT ACACCTCGCC
22801 CTGCTCGGGC ATCAGCTGGA AGCGGACTT CAGGTCGCTC TCCACGCGGT ACCGGTCCAT
22861 CAGCAGCGTC ATCACTTCCA TGCCCTTCTC CCAGGCCGAA ACGATCGGCA GGCTCAGGGG
22921 GTTCTTCACC GCCATTGTCA TCTTAGTTCG CGCCGCCGAG GTCAGGGGGT CGTTCTCGTC
22981 CAGGGTCTCA AACACTCGCT TGCCGTCTTT CTCGATGATG CGCACGGGGG GAAAGCTGAA
23041 GCCCAGGGCC GCCAGCTCCT CCTCGGCCCT CTTTCGTTCC TCGCTGTCTT GGCTGATGTC
23101 TTGCAAAGGC ACATGCTTGG TCTTGCGGGG TTTCTTTTTC GCGGCGAGAG GCGGCGGCGA
23161 TGTGCTGGGA GAGCGCGAGT TCTCGTTTAC CACGACTATT TCTTCTTCTT GGCCGTCTGTC
23221 CGAGCCACAG CGGCGGTAGG CATGCCCTCT CTGGGGCAGA GGCGGAGGCG ACGGGCTCTC
23281 GCGGTTCCGC GGGCGGCTGG CAGAGCCCTT TCCGCGTTCG GGGGTGCGCT CCTGGCGGCG
23341 CTGCTCTGAC TGACTTCCTC CGCGGCCGCG CATTGTGTTT TCCTAGGGAG CAACAACAAG
23401 CATGGAGACT CAGCCATCGT CGCCAACATC GCCATCTGCC CCCGCCGCA CCGCCGACGA
23461 GAACCAGCAG CAGAATGAAA GCTTAACCGC CCCGCCGCC AGCCCCACCT CCGACGCGCG
23521 GGCCCCAGAC ATGCAAGAGA TGGAGGAATC CATCGAGATT GACCTGGGCT ACGTGACGCC
23581 CGCGGAGCAC GAGGAGGAGC TGGCAGCGCG CTTTTCAGCC CCGGAAGAGA ACCACCAAGA
23641 GCAGCCAGAG CAGGAAGCAG AGAACGAGCA GAACCAGGCT GGGCAGGAGC ATGGCGACTA
23701 CCTGAGCGGG GCAGAGGACG TGCTCATCAA GCATCTGGCC CGCCAATGCA TCATCGTCAA
23761 GGACGCGCTG CTCGACCGCG CCGAGGTGCC CCTCAGCGTG GCGGAGCTCA GCCGCGCCTA
23821 CGAGCGCAAC CTCTTCTCGC CGCGCGTGCC CCCCAGCGC CAGCCCAACG GCACCTGTGA
23881 GCCCCAACCG CGCCTCAACT TCTACCCGGT CTTGCGGGTG CCCGAGGCCC TGGCCACCTA
23941 CCACCTCTTT TTCAAGAACC AAAGGATCCC CGTCTCCTGC CGCGCCAACC GCACCCGCGC
24001 CGACGCCCTG CTCAACCTGG GCCCCGGCGC CCGCCTACCT GATATCACCT CCTTGGAAGA
24061 GGTTCCCAAG ATCTTCGAGG GTCTGGGCAG CGACGAGACT CGGGCCGCGA ACGTCTGCA
24121 AGGAAGCGGA GAGGAGCATG AGCACCATG CGCCCTGGTG GAGTTGGAAG GCGACAACGC
24181 GCGCCTGGCG GTCTTCAAGC GCACGGTCGA CTTGACCCAC TTCGCTTACC CGGCGCTCAA
24241 CCTGCCCCC AAGGTCATGA GCGCCGTCAT GGACCAGGTG CTCATCAAGC GCGCCTCGCC
24301 CCTCTCGGAG GAGGAGATGC AGGACCCCGA GAGTTCGGAC GAGGGCAAGC CCGTGGTCAG
24361 CGACGAGCAG CTGGCGCGCT GGCTGGGAGC GAGTANACC CCCCAGAGCC TGGAAGAGCG
24421 GCGCAAGCTC ATGATGGCCG TGGTCTTGGT GACCGTGGAG CTGGAGTGTG TGCGCCGCTT
24481 CTTTGCCGAC GCGGAGACCC TGCGCAAGGT CGAGGAGAAC CTGCACTACC TCTTCAGGCA
24541 CCGGTTCTGT CGCCAGGCTT GCAAGATCTC CAACGTGGAG CTGACCAACC TGGTCTCCTA
24601 CATGGGCATC CTGCACGAGA ACCGCTGGG GCAAAACGTG CTGCACACCA CCCTGCGCGG
24661 GGAGGCCCGC CGCGACTACA TCCGCGACTG CGTCTACCTG TACCTCTGCC ACACCTGGCA
24721 GACGGGCATG GCGGTGTGGC AGCAGTGCCT GGAGGAGCAG AACCTGAAAG AGCTCTGCAA
24781 GCTCCTGCAG AAGAACCTCA AGGCCCTGTG GACCGGGTTC GACGAGCGTA CCACCGCCTC
24841 GGACCTGGCC GACCTCATCT TCCCGAGCG CCGCGGCTG ACGCTGCGCA ACGGGCTGCC
24901 CGACTTTATG AGCCAAAGCA TGTTGCAAAA CTTTCGCTCT TTCATCCTCG AACGCTCCGG
24961 GATCCTGCCC GCCACCTGCT CCGCGCTGCC CTCGGACTTC GTGCCGCTGA CTTCCGCGA
25021 GTGCCCCCG CCGCTCTGGA GCCACTGCTA CTTGCTGCGC CTGGCCAACT ACCTGGCCTA
25081 CCACTCGGAC GTGATCGAGG ACGTCAGCG CGAGGGTCTG CTGGAGTGCC ACTGCCGCTG
25141 CAACCTCTGC ACGCCGCACC GCTCCCTGGC CTGCAACCCC CAGCTGCTGA GCGAGACCCA
25201 GATCATCGGC ACCTTCGAGT TGCAAGGCCC CGCGACGCG GAGGGCAAGG GGGGTCTGAA

Fig. 7G

25261 ACTCACCCCG GGGCTGTGGA CCTCGGCCTA CTTGCGCAAG TTCGTGCCCC AGGACTACCA
25321 TCCCTTCGAG ATCAGGTTCT ACGAGGACCA ATCCCAGCCG CCCAAGGCCG AGCTGTCGGC
25381 CTGCGTCATC ACCCAGGGGG CCATCCTGGC CCAATTGCAA GCCATCCAGA AATCCCGCCA
25441 AGAATTCTCTG CTGAAAAAGG GCCACGGGGT CTACTTGGAC CCCAGACCG GAGAGGAGCT
25501 CAACCCAGC TTCCCCCAGG ATGCCCCGAG GAAGCAGCAA GAAGCTGAAA GTGGAGCTGC
25561 CGCCGCCGGA GGATTTGGAG GAAGACTGGG AGAGCAGTCA GGCAGAGGAG GAGGAGATGG
25621 AAGACTGGGA CAGCACTCAG GCAGAGGAG ACAGCCTGCA AGACAGTCTG GAGGAGGAAG
25681 ACGAGGTGGA GGAGGCAGAG GAAGAAGCAG CCGCCGCCAG ACCGTCTGTC TCGGCGGAGA
25741 AAGCAAGCAG CACGGATACC ATCTCCGCTC CGGGTCGGGG TCGCGGCGGC CGGGCCCACA
25801 GTAGGTGGGA CGAGACCGGG CGTTCCGAA CCCCACCACC CAGACCGGTA AGAAGGAGCG
25861 GCAGGGATAC AAGTCCTGGC GGGGGCACA AAACGCCATC GTCTCTTGCT TGCAAGCCTG
25921 CGGGGGCAAC ATCTCCTTCA CCCGGCGCTA CCTGCTCTTT CACCGCGGGG TGAACCTCCC
25981 CCGCAACATC TTGCATTACT ACCGTCACCT CCACAGCCCC TACTACTGTT TCCAAGAAGA
26041 GGCAGAAACC CAGCAGCAGC AGAAGAGCAG AGCTAGAAAA TCCACAGCGG
26101 CGGCAGGTGG ACTGAGGATC GCGGCGAAGC AGCCGGCGCA GACCCGGGAG CTGAGGAACC
26161 GGATCTTTCC CACCTCTAT GCCATCTTCC AGCAGAGTCG GGGGCAGGAG CAGGAGCTGA
26221 AAGTCAAGAA CCGTTCTCTG CGCTCGCTCA CCCGCACTTG TCTGTATCAC AAGAGCGAAG
26281 ACCAATTCA GCGCACTCTC GAGGACGCCG AGGCTCTCTT CAACAAGTAC TGCGCGCTCA
26341 CTCTTAAAGA GTAGCCCGCG CCCGCCACA CACGGAAGAA GCGCGGAATT ACCTCACCAC
26401 CTGCGCCCTT CGCCGACCA TCATGAGCAA AGAGATTCCC ACGCCTTACA TGTGGAGCTA
26461 CCAGCCCCAG ATGGGCCTGG CCGCCGGCGC CGCCAGGAC TACTCCACCC GCATGAAGT
26521 GCTCAGTGCC GGGCCCGCGA TGATCTCAGC GGTGAATGAC ATCCGCGCCC ACCGAAACCA
26581 GATACTCCTA GAACAGTCAG CGATCACCGC CACGCCCCCG CATCACCTTA ATCCGCGTAA
26641 TTGGCCCCGC GCCCTGGTGT ACCAGGAAAT TCCCCAGCCC ACGACCGTAC TACTTCCGCG
26701 AGACGCCCGC GCGGAAGTCC AGCTGACTAA CTCAGGTGTC CAGCTGGCCG GCGGCGCCGC
26761 CCTGTGTCGT CACCGCCCCG CTCAGGGTAT AAAGCGGCTG GTGATCCGAG GCAGAGGCAC
26821 ACAGCTCAAC GACGAGGTGG TGAGCTCTTC GCTGGGTCTG CGACCTGACG GAGTCTTCCA
26881 ACTCGCCGGA TCGGGGAGAT CTTCTTTCAC GCCTCGTCAG GCCGTCTGA CTTTGGAGAG
26941 TTCGTCTCG CAGCCCCGCT CGGGCGGCAT CGGCACTCTC CAGTTCGTGG AGGAGTTCAC
27001 TCCCTCGGTN TACTTCAACC CTTTCTCCGG CTCCCCCGGC CACTACCCGG ACAGGTTTCC
27061 CCCGAACCTC GACGCCATCA GCGAGTCGGT GGACGGCTAC GATTGAATGT CCCATGGTGG
27121 CGCAGCTGAC CTAGCTCGGC TTCGACACCT GGACACTGC CGCCGCTTCC GCTGCTTCGC
27181 TCGGGATCTC GCGGAGTTTG CCTACTTTGA GCTGCCCGAG GAGCACCTC AGGGCCAGC
27241 CCACGGAGTG CGGATCATCG TCGAAGGGGG CCTCGACTCC CACCTGCTTC TGTACTGCAT
27301 CCAGCGACCG ATCCTGGTCC AGCGCGAACA AGGACAGACC CTTCTTACTT AGTATAATAA
27361 CTGCAACCAC CCCGGCCTGC ATGAAAGTCT TTGTTGTCTG CTGTGTACTG AACCGGTTCC
27421 AAGTGAGAT CAGCGACTAC TCCGGACTCG ATTGTGGTGT TCCTGCTATC AACCCTTCCC
27481 TGTCTTTCAC CGGGAACGAG ACCGAGCTCC AGCTCCAGTG TAAGCCCCAC AAGAAGTACC
27541 TCACCTGGCT GTTCCAGGGC TCCCCGATCG CCGTTGTCAA CCACTGCGAC AACGACGGAG
27601 TCCTGCTGAG CGGCCCTGCC AACCTTACTT TTTCCACCCG CAGAAGCAAG CTCCAGCTCT
27661 TCCAACCCCTT CCTCCCCGGG ACCATACAGT GCGTCTCAGG ACCCTGCCAT CACACCTTCC
27721 ACCTGATCCC GAATACCACA CGCGCTACTA CAACCAAAC ACACCAACAC
27781 GCCACCGTCG CGACCTTTCC TCTGAATCTA ATACCACTAC CGGAGGTGAG CTCCGAGGTC
27841 GACCAACCTC TGGGATTTAC TACGGCCCCC GGGAGGTGGT GGGGTTAATA GCGCTAGGCC
27901 TAGTTGCGGG TGGGCTTTTG GTTCTCTGCT ACCTATACCT CCCTTGCTGT TCGTACTTAG
27961 TGGTGCTGTG TTGCTGTTT AAGAAATGGG GAAGATCACC CTAGTGAGCT GCGGTGCGCT
28021 GGTGGCGGTG TTGCTTTCGA TTGTGGGACT GGGCGGCGCG GCTGTAGTGA AGGAGAAGGC
28081 CGATCCCTGC TTGCATTCA ATCCCAACA ATGCCAGCTG AGTTTTCAGC CCGATGGCAA
28141 TCGGTGCGCG GTACTGATCA AGTGCAGATG GGAATGCGAG AACGTGAGAA TCGAGTACAA
28201 TAACAAGACT CGGAACAATA CTCTCGCGTC CGTGTGGCAG CCCGGGGACC CCGATGGTA
28261 CACCGTCTCT GTCCCCGGTG CTGACGGCTC CCCGCGCACC GTGAATAATA CTTTCACTTT
28321 TGCGCACATG TGCAACACGG TCATGTGGAT GAGCAAGCAG TACGATATGT GGGCCCCAC
28381 GAAGGAGAAC ATCGTGGTCT TCTCCATCGC TTACAGCCTG TGCACGGCGC TAATCACCAG
28441 TATCGTGTGC CTGAGCATTC ACATGCTCAT CGCTATTTCG CCCAGAAATA ATGCCGAGAA
28501 AGAGAAACAG CCATAACACG TTTTTCACA CACCTTGTTT TTACAGACAA TGCGTCTGTT
28561 AAATTTTTTA AACATTGTGC TCAGTATGTC TTATGCCTCT GGTATGCAA ACATACAGAA
28621 AACCTTTTAT GTAGGATCTG ATGGTACACT AGAGNGTACC CAATCACAAG CCAAGGTGTC
28681 ATGGTATTTT TATAGAACCA AACTGATCC AGTTAACTT TGTAAGGGTG AATTGCGCGC
28741 TACACATAAA ACTCCACTTA CATTTAGTTG CAGCAATAAT AATCTTACAC TTTTTCAT
28801 TACAAACAA TATACTGGTA CTTATTACAG TACAACTTT CATAAGGAC AAGATAAATA

Fig. 7H


```

28861 TTATACTGTT AAGGTAGAAA ATCCTACCAC TCCTAGAACT ACCACCACCA CCACTACTGC
28921 AAAGCCCACT GTGAAAACCTA CAACTAGGAC CACCACAACCT ACAGAAACCA CCACCAGCAC
28981 AACACTTGCT GCAACTACAC ACACACACAC TAAGCTAACC TTACAGACCA CTAATGATTT
29041 GATCGCCCTG CTGCAAAAGG GGGATAACAG CACCACTTCC AATGAGGAGA TACCCAAATC
29101 CATGATTGGC ATTATGTGTG CTGTAGTGGT GTGCATGTTG ATCATCGCCT TGTGCATGGT
29161 GTACTATGCC TTCTGCTACA GAAAGCACAG ACTGAACGAC AAGCTGGAAC ACTTACTAAG
29221 TGTGAATTT TAATTTTTTA GAACCATGAA GATCCTAGGC CTTTTTAGTT TTTCTATCAT
29281 TACCTCTGCT CTTTGTGAAT CAGTGGATAG AGATGTTACT ATTACCACTG GTTCTAATTA
29341 TACACTGAAA GGGCCACCCT CAGGTATGCT TTCGTGGTAT TGCTATTTTG GAACTGACAC
29401 TGATCAAACT GAATTATGCA ATTTTCAAAA AGGCAAAACC TCAAACCTA AAATCTCTAA
29461 TTATCAATGC AATGGCACTG ATCTGATACT ACTCAATGTC ACGAAAGCAT ATGGTGGCAG
29521 TTATTATTGC CCTGGACAAA AACTGAAGA AATGATTTT TACAAAGTGG AAGTGGTTGA
29581 TCCCCTACA CCACCCACCA CCACAACAT TCATACCACA CACACAGAAC AAACACCAGA
29641 GGCAACAGAA GCAGAGTTGG CTTTCCAGGT TCACGGAGAT TCCTTTGCTG TCAATACCCC
29701 TACACCCGAT CAGCGGTGTC CGGGGCCGCT AGTCAGCGGC ATTGTCGGTG TGCTTTCCGG
29761 ATTAGCAGTC ATAATCATCT GCATGTTTCT TTTTGCCTTG TGCTATAGAA GGCTTTACCG
29821 ACAAAAATCA GACCCACTGC TGAACCTCTA TGTTTAATTT TTTCCAGAGC CATGAAGGCA
29881 GTTAGCGCTC TAGTTTTTTG TTCTTTGATT GGCATTGTTT TTAATAGTAA AATTACCAGA
29941 GTTAGCTTTA TTAAACATGT TAATGTAAC GAAGGAGATA ACATCACACT AGCAGGTGTA
30001 GAAGGTGCTC AAAACACCAC CTGGACAAAA TACCATCTAG GATGGAGAGA TATTTGCACC
30061 TGAATGTAA CTTATTATTG CATAGGAGTT AATCTTACCA TTGTTAACGC TAACCAATCT
30121 CAGAATGGGT TAATTAAAGG ACAGAGTGTT AGTGTGACCA GTGATGGGTA CTATACCCAG
30181 CATAGTTTTA ACTACAACAT TACTGTCATA CCACTGCCTA CGCCTAGCCC ACCTAGCACT
30241 ACCACACAGA CAACCACATA CAGTACATCA AATCAGCCTA CCACCACTAC AGCAGCAGAG
30301 GTTGCCAGCT CGTCTGGGGT CCGAGTGGCA TTTTGTATGT TGGCCCCATC TAGCAGTCCC
30361 ACTGCTAGTA CCAATGAGCA GACTACTGAA TTTTGTCCA CTGTCGAGAG CCACACCACA
30421 GCTACCTCCA GTGCCTTCTC TAGCACCGCC AATCTCTCCT CGCTTTCTCT TACACCAATC
30481 AGCCCCGCTA CTACTCCTAG CCCCCTCCT CTTCCCACTC CCCTGAAGCA AACAGACGGC
30541 GGCATGCAAT GGCAGATCAC CCTGCTCAT GTGATCGGGT TGGTCATCCT GGCGTGTGTG
30601 CTCTACTACA TCTTCTGCCG CCGCATTCAC AACGCGCACC GCAAGCCGGC CTACAAGCCC
30661 ATCGTTATCG GGCAGCCGGA GCCGCTCCG GTGGAAGGGG GTCTAAGGAA TCTTCTCTTC
30721 TCTTTTACAG TATGGTGATT GAANTATGAT TCCTAGACAA TTCTTGATCA CTATTCTTAT
30781 CTGCCTCCTC CAAGTCTGTG CCACCTCGC TCTGGTGGCC AACGCCAGTC CAGACTGTAT
30841 TGGGCCCTTC GCCTCCTACG TGCTCTTTGC CTTCTGTCACC TGCATCTGCT GCTGTAGCAT
30901 AGTCTGCCTG CTTATCACCT TCTTCCAGTT CATTGACTGG ATCTTTGTGC GCATCGCCTA
30961 CCTGCGCCAC CACCCCCAGT ACCGCGACCA GCGAGTGGCG CAGCTGCTCA GGCTCCTCTG
31021 ATAAGCATGC GGGCTCTGCT ACTTNTCGCG CTTCTGCTGT TAGTGTCTCC CCGTCCCGTC
31081 GACCCCGGT CCCCCACTCA GTCCCCCGAG GAGGTTTCGCA AATGCAAAT CCAAGAACCC
31141 TGGAAATTC TCAAATGCTA CCGCCAAAAA TCAGACATGC ATCCCAGCTG GATCATGATC
31201 ATTGGGATCG TGAACATTCT GGCCTGCACC CTCATCTCCT TTGTGATTTA CCCCTGCTTT
31261 GACTTTGGTT GGAACTCGCC AGAGGCGCTC TATCTCCCGC CTGAACCTGA CACACCACCA
31321 CAGCATCAAC CTCAGGCACA CGCACTACCA CCACCACAGC CTAGGCCACA ATACATGCCC
31381 ATATTAGACT ATGAGGCCGA GCCACAGCGA CCCATGCTCC CCGCTATTAG TTAATTCAAT
31441 CTAACCGGCG GAGATGACTG ACCCATGGC CAATAACAAC GTCAACGACC TTCTCCTGGA
31501 CATGGACGGC CGCGCTCGG AGCAGCGACT CGCCCAACTT CGCATTCGTC AGCAGCAGGA
31561 GAGAGCCGTC AAGGAGCTGC AGGACGGCAT AGCCATCCAC CAGTGCAAGA GAGGCATCTT
31621 CTGCCTGGTG AAACAGGCCA AGATCTCCTA CGAGGTCACC CAGACCGACC ATCGCCTCTC
31681 CTACGAGCTC CTGCAGCAGC GCCAGAAGTT CACCTGCCTG GTCGGAGTCA ACCCCATCGT
31741 CATCACCCAG CCAGCAGTCG GGCAGATACA AGGGGTGCAT CCACTGCTCC TGCAGTCCC
31801 CCGACTGCGT CCACACTCTG ATCAAGACCC TCTGCGGCCCT CCGCGACCTC CTCCCCATGA
31861 ACTAATCACC CCCTTATCCA GTGAATAAAA GATCATATTG ATGATGATTT AAATAAAAAA
31921 AATAATCATT TGATTTGAAA TAAAGATACA ATCATATTGA TGATTTGAGT TTAACAAAAA
31981 TAAAGAATCA CTTACTTGAA ATCTGATACC AGGTCTCTGT CCATGTTTTT TGCCAACACC
32041 ACCTCACTCC CCTCTTCCCA GCTCTGGTAC TGCAGGCCCC GGCGGGCTGC AAACCTCTCT
32101 CACACGCTGA AGGGGATGTC AAATTCCCTC TGTCCTCAA TCTTCATTTT ATCTTCTATC
32161 AGATGTCCAA AAAGCGCGTC CGGGTGGATG ATGACTTCGA CCCCCTCTAC CCCTACGATG
32221 CAGACAACGC ACCGACCGTG CCCTTCATCA ACCCCCCCTT CGTCTCTTCA GATGGATTCC
32281 AAGAGAAGCC CCTGGGGGTG TTGTCCCTGC GACTGGCTGA CCCCCTCACC ACCAAGAACG
32341 GGGAAATCAC CCTCAAGCTG GGAGAGGGGG TGGACCTCGA CTCGTCGGGA AAACATCATCT
32401 CCAACACGGC CACCAAGGCC GCCGCCCCCTC TCAGTATTTT AAACAACACC ATTTCCCTTA

```

Fig. 7I

32461 AAACGTGCTGC CCCTTTCTAC AACAAACAATG GAACTTTAAG CCTCAATGTC TCCACACCAT
32521 TAGCAGTATT TCCCACATTT AACACTTTAG GCATAAGTCT TGGAAACGGT CTTCAGACTT
32581 CAAATAAGTT GTTGACTGTA CAACTAACTC ATCCTCTTAC ATTCAAGCTCA AATAGCATCA
32641 CAGTAAAAAC AGACAAAGGG CTATATATTA ACTCCAGTGG AAACAGAGGA CTGAGGCTA
32701 ATATAAGCCT AAAAAGAGGA CTAGTTPTTG ACGGTAATGC TATTGCAACA TATATTGGAA
32761 ATGGCTTAGA CTATGGATCT TATGATAGTG ATGGAAAAAC AAGACCCGTA ATTACCAAAA
32821 TTGGAGCAGG ATTAATATTTT GATGCTAACA AAGCAATAGC TGTCAAAC TA GGCACAGGTT
32881 TAAGTTTTGA CTCCGCTGGT GCCTTGACAG CTGGAAACAA ACAGGATGAC AAGCTAACAC
32941 TTTGGACTAC CCCTGACCCA AGCCCTAATT GTCAATTACT TTCAGACAGA GATGCCAAAT
33001 TTACTCTCTG TCTTACAAAA TGCGGTAGTC AAATACTAGG CACTGTGGCA GTGGCGGCTG
33061 TTACTGTAGG ATCAGCACTA AATCCAATTA ATGACACAGT CAAAAGCGCC ATAGTTTTCC
33121 TTAGATTTGA TTCCGATGGT GTACTCATGT CAAACTCATC AATGGTAGGT GATTACTGGA
33181 ACTTTAGGGA GGGACAGACC ACTCAAAGTG TAGCCTATAC AAATGCTGTG GGATTTCATGC
33241 CAAATATAGG TGCATATCCA AAAACCCAAA GTAAACACACC TAAAAATAGC ATAGTCAGTC
33301 AGGTATATTT AACTGGAGAA ACTACTATGC CAATGACACT AACCATAACT TTCAATGGCA
33361 CTGATGAAAA AGACACAACC CCAGTTAGCA CCTACTCTAT GACTTTTACA TGGCAGTGGGA
33421 CTGGAGACTA TAAGGACAAA AATATTACCT TTGCTACCAA CTCATTCTCT TTTTCTTACA
33481 TCGCCAGGA ATAATCCAC CCAGCAAGCC AACCCTTTT CCCACCACCT TTGTCTATAT
33541 GGAAACTCTG AAACAGAAAA ATAAAGTTCA AGTGTTTTAT TGAATCAACA GTTTTACAGG
33601 ACTCGAGCAG TTATTTTCC TCCACCCTTC CAGGACATGG AATACACCAC CCTCTCCCCC
33661 CGCACAGCCT TGAACATCTG AATGCCATTG GTGATGGACA TGCTTTTGGT CTCCACGTTT
33721 CACACAGTTT CAGAGCGAGC CAGTCTCGGA TCGGTCAGGG AGATGAAACC CTCCGGGCAC
33781 TCCCGCATCT GCACCTCACA GCTCAACAGC TGAGGATTGT CCTCGGTGGT CGGGATCAGC
33841 GTTATCTGGA AGAAGCAGAA GAGCGCGGGT GGAATCATA GTCCGCGAAC GGGATCGGCC
33901 GGTGGTGTG CATCAGGCCC CGCAGCAGTC GCTGCCGCCG CCGCTCCGTC AAGCTGCTGC
33961 TCAGGGGGTT CGGGTCCAGG GACTCCCTCA GCATGATGCC CACGGCCCTC AGCATCAGTC
34021 GTCCTGGTGG CCGGGCGCAG CAGCGCATGC GAATCTCGCT CAGGTCACTG CAGTACGTGC
34081 AACACAGGAC CACCAGGTG TTCAACAGT CATAGTTCAA CACGCTCCAG CCGAACTCA
34141 TCGCGGGAAG GATGCTACCC ACGTGGCCGT CGTACCAGAT CCTCAGGTAA ATCAAGTGGC
34201 GCTCCCTCCA GAAGACGCTG CCCATGTACA TGATCTCCTT GGGCATGTGG CGGTTACCA
34261 CCTCCCGTA CCACATCACC CTCTGGTTGA ACATGCAGCC CCGGATGATC CTGCGGAACC
34321 ACAGGGCCAG CACCGCCCCG CCCGCCATGC AGCGAAGAGA CCGCGGATCC CGGCAATGAC
34381 AATGGAGGAC CCACCGCTCG TACCCGTGGA TCATCTGGGA GCTGAACAAG TCTATGTTGG
34441 CACAGCAGC GCATATGCTC ATGCATCTCT TCAGCACTCT CAGCTCCTCG GGGGTCAAAA
34501 CCATATCCCA GGGCAGGGG AACTCTTGCA GGACAGCGAA CCGCGCAGAA CAGGGCAATC
34561 CTCGCACATA ACTTACATTG TGCATGTCGA GGGTATCGCA ATCAGGCAGC ACCGGGTGAT
34621 CCTCCACCAG AGAAGCGCGG GTCTCGGTCT CCTCACAGCG TGGTAAGGGG GCCGGCCGAT
34681 ACGGGTGATG GCGGGACGCG GCTGATCGTG TTCTCGACCG TGTCATGATG CAGTTGCTTT
34741 CGGACATTTT CGTACTTGCT GTAGCAGAAC CTGGTCCGGG CGCTGCACAC CGATCGCCGG
34801 CGGCGGTCTC GGCGCTTGGA ACGCTCGGTG TTAAAGTTGT AAAACAGCCA CTCTCTCAGA
34861 CCGTGCAGCA GATCTAGGCA CTCAGGAGTG ATGAAGATCC CATCATGCCT GATAGCTCTG
34921 ATCATATCGA CCACCGTGGG ATGGGCCAGG CCCAGCCAGA TGATGCAATT TTGTTGGGTT
34981 TCGGTGACGG CGGGGGAGGG AAGAACAGGA AGAACCATGA TTAACTTTTA ATCCAAACGG
35041 TCTCGGAGCA CTTCAAATG AAGGTCACGG AGATGGCACC TCTCGCCCCG GCTGTGTTGG
35101 TGGAAAATAA CAGCCAGGTC AAAGGTGATA CGGTTCTCGA GATGTTCCAC GGTGGCTTCC
35161 AGCAAAGCCT CCACGCGCAC ATCAGAAACA AGACAATAGC GAAAGCGGGA GGGTCTCTTA
35221 ATTCCTCAAC CATCATGTTA CACTCCTGCA CCATCCCCAG ATAATTTTCA TTTTCCAGC
35281 CTTGAATGAT TCGAACTAGT TCCTGAGGTA AATCCAAGCC AGCCATGATA AAAAGCTCGC
35341 GCAGAGCACC CTCCACCGGC ATTCTTAAGC ACACCCTCAT AATTCCAAGA TATTCTGCTC
35401 CTGGTTTACC TGCAGCAGAT TGACAAGCGG AATATCAAAA TCTCTGCCGC GATCCCTGAG
35461 CTCCTCCCTC AGCAATAACT GTAAGTACTC TTTCATATCG TCTCCGAAAT TTTTAGCCAT
35521 AGGACCCCCA GGAATAAGAG AAGGGCAAGC CACATTACAG ATAAACCGAA GTCCCCCCCCA
35581 GTGAGCATTTG CCAAATGTAA GATTGAAATA AGCATGCTGG CTAGACCCGG TGATATCTTC
35641 CAGATAACTG GACAGAAAAAT CGGGTAAGCA ATTTTAAAGA AAATCAACAA AAGAAAAATC
35701 TTCCAGGTGC ACGTTTAGGG CCTCGGGAAC AACGATGGAG TAAGTGCAAG GGGTGCGTTT
35761 CAGCATGGTT AGTTAGCTGA TCTGTAAAAA AACAAAAAAT AAAACATTAA ACCATGCTAG
35821 CCTGGCGAAC AGGTGGGTAA ATCGTTCTCT CCAGCACCAG GCAGGCCACG GGGTCTCCGG
35881 CGCGACCCCT GTAAAAATTG TCGCTATGAT TGAAAACCAT CACAGAGAGA CGTTCCCGGT
35941 GGCCGGCGTG AATGATTCTGA GAAGAAGCAT ACACCCCCCG GAACATTGGA GTCCGTGAGT
36001 GAAAAAAGC GGCCGAGGAA GCAATGAGGC ACTACAACGC TCACTCTCAA GTCCAGCAAA

Fig. 7J

ITR0048PV

SEQ ID NO: 3

59/153

36061 GCGATGCCAT GCGGATGAAG CACAAAATTT TCAGGTGCGT AAAAAATGTA ATTACTCCCC
36121 TCCTGCACAG GCAGCGAAGC TCCCGATCCC TCCAGATACA CATACAAAGC CTCAGCGTCC
36181 ATAGCTTACC GAGCGGCAGC AGCAGCGGCA CACAACAGGC GCAAGAGTCA GAGAAAAGAC
36241 TGAGCTCTAA CCTGTCCGCC CGCTCTCTGC TCAATATATA GCCCCAGATC TACTACTGACG
36301 TAAAGGCCAA AGTCTAAAAA TACCCGCCAA ATAATCACAC ACGCCCAGCA CACGCCCAGA
36361 AACCGGTGAC AACTCAGAA AAATACGCGC ACTTCCTCAA ACGGCCAAAC TGCCGTCATT
36421 TCCGGGTTCC CACGCTACGT CATCAAAACA CGACTTTCAA ATTCCGTCGA CCGTTAAAAA
36481 CATCACCCGC CCCGCCCTA ACGGTCGCCG CTCCCGCAGC CAATCACCTT CCTCCCTCCC
36541 CAAATTCAAA CAGCTCATTT GCATATTAAC GCGCACCAA AGTTTGAGGT ATATTATTGA
36601 TGATGG

Fig. 7K

SEQ ID No: 4

60/153

```

1  CATCATCAAT AATATACCTC AAACTTTTGG TGC GCGTTAA TATGCAAATG AGCTGTTTGA
61 ATTTGGGGGAG GGAGGAAGGT GATTGGCCGA GAGACGGGCG ACCGTTAGGG GCGGGGCGGG
121 TGACGTTTTG ATGACGTGGC CGTGAGGCGG AGCCGGTTTG CAAGTTCTCG TGGGAAAAGT
181 GACGTCAAAC GAGGTGTGGT TTGAACACGG AAATACTCAA TTTTCCCGCG CTCCTGTACA
241 GGAAATGAGG TGTTTCTGGG CGGATGCAAG TGAAAACGGG CCATTTTCGC GCGAAAAC TG
301 AATGAGGAAG TGAAAATCTG AGTAATTTCTG CGTTTATGGC AGGGAGGAGT ATTTGCCGAG
361 GGCCGAGTAG ACTTTGACCG ATTACGTGGG GGTTCGATT ACCGTATTTT TCACCTAAAT
421 TTCCGCGTAC GGTGTCAAAG TCCGGTGT TTACGTAGGC GTCAGCTGAT CGCCAGGGTA
481 TTTAAACCTG CGCTCTCTAG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGTTTCT
541 CCTCCGCGCC GCGAGTCAGA TCTACACTTT GAAAGATGAG GCACCTGAGA GACCTGCCCG
601 GTAATGTTTT CCTGGCTACT GGAACGAGA TTCTGGAATT GGTGGTGGAC GCCATGTGG
661 GTGGCGACCC TCCTGAGCCC CTTACCCCAT TTGAGGCGCC TTCGCTGTAC GATTTGTATG
721 ATCTGGAGGT GGATGTGCCC GAGAACGACC CCAACGAGGA GGCGGTGAAT GATTTGTTTA
781 GCGATGCCGC GCTGCTGGCT GCCGAGCAGG CTAATACGGA CTCTGGCTCA GACAGCGATT
841 CCTCTCTCCA TACCCCGAGA CCCGGCAGAG GTGAGAAAAA GATCCCCGAG CTTAAAGGGG
901 AAGAGCTCGA CCTGCGCTGC TATGAGGAAT GCTTGCCCTCC GAGCGATGAT GAGGAGGACG
961 AGGAGGCGAT TCGAGCTGCA TCGAACCAGG GAGTGAAAGC TCGGGCGGAA AGCTTTAGCC
1021 TGGACTGTCC TACTCTGCCC GGACACGGCT GTAAGTCTTG TGAATTTTCA CGCATGATA
1081 CTGGAGATAA GAATGTGATG TGTGCCCTGT GCTATATGAG AGCTTACAAC CATTGTGTTT
1141 ACAGTAAGTG TGATTAACCT TAGTTGGGAA GGCAGAGGGT GACTGGGTGC TGACTGGTTT
1201 ATTTATGTAT ATGTTTTTTT ATGTGTAGGT CCCGTCTCTG ACGTAGATGA GACCCCACT
1261 TCAGAGTGCA TTTTCATCACC CCCAGAAAT GGCGAGGAAC CGCCCAAGA TATTATTCAT
1321 AGACCAGTTG CAGTGAGAGT CACCGGGCGG AGAGCAGCTG TGGAGAGTTT GGATGACTTG
1381 CTACAGGGTG GGGATGAACC TTTGGACTTG TGTACCCGGA AACGCCCGAG GCACTAAGTG
1441 CCACACATGT GTGTTTACTT AAGGTGATGT CAGTATTTAT AGGGTGTGGA GTGCAATAAA
1501 ATCCGTGTTG ACTTTAAGTG CGTGGTTTAT GACTCAGGGG TGGGGACTGT GGGTATATAA
1561 GCAGGTGCAG ACCTGTGTGG TCAGTTCAGA GCAGGACTCA TGGAGATCTG GACGGTCTTG
1621 GAAGACTTTC ACCAGACTAG ACAGCTGCTA GAGAACTCAT CGGAGGGGGT CTCTTACCTG
1681 TGGAGATTCT GCTTCGGTGG GCCTCTAGCT AAGCTAGTCT ATAGGGCCAA ACAGGATTAT
1741 AAGGATCAAT TTGAGGATAT TTTGAGAGAG TGTCTGGTA TTTTGTACTC TCTCAACTTG
1801 GGCCATCAGT CTCACCTTAA CCAGAGTATT CTGAGAGCCC TTGACTTTTC TACTCCTGGC
1861 AGAACTACCG CCGCGGTAGC CTTTTTTGCC TTTATCCTTG ACAAATGGAG TCAAGAAACC
1921 CATTTTCAGCA GGGATTACCG TCTGGACTGC TTAGCAGTAG CTTTGTGGAG AACATGGAGG
1981 TGCCAGCGCC TGAATGCAAT CTCCGGCTAC TTGCCAGTAC AGCCGGTAGA CACGCTGAGG
2041 ATCCTGAGTC TCCAGTCACC CCAGGAACAC CAACGCCGCC AGCAGCCGCA GCAGGAGCAG
2101 CAGCAAGAGG AGGAGGAGGA TCGAGAAGAG AACCCGAGAG CCGGTCTGGA CCTCCGGTG
2161 GCGGAGGAGG AGGAGTAGCT GACTTGTTC CCGAGCTGCG CCGGTGCTG ACTAGGTCTT
2221 CCAGTGAGCG GGAGAGGGGG ATTAAGCGGG AGAGGCATGA GGAGACTAGC CACAGAAGTG
2281 AACTGACTGT CAGTCTGATG AGCCGACGGC GCCAGAATC GGTGTGGTGG CATGAGGTTT
2341 AGTCGCAGGG GATAGATGAG GTCTCGGTGA TGCATGAGAA ATATTCCCTG GAACAAGTCA
2401 AGACTTGTG GTTGGAGCCT GAGGATGATT GGGAGGTAGC CATCAGGAAT TATGCCAAGC
2461 TGGCTCTGAA GCCAGACAAG AAGTACAAGA TTACCAAAC GATTAATATC AGAAATTCCT
2521 GCTACATTTT AGGGAATGGG GCCGAGGTGG AGATCAGTAC CCAGGAGAGG GTGGCCTTCA
2581 GATGTTGTAT GATGAATATG TACCCGGGGG TGGTGGGCAT GGAGGGAGTC ACCTTTATGA
2641 AC CGGAGGTT CAGGGGTGAT GGGTATAATG GGGTGGTCTT TATGGCCAAC ACCAAGCTGA
2701 CAGTGCACGG ATGCTCCTTC TTTGGGTTCA ATAACATGTG CATCGAGGCC TGGGGCAGTG
2761 TTTCACTGAG GGGATGCAGC TTTTCAGCCA ACTGGATGGG GGTGCTGGGC AGAACCAAGA
2821 GCAAGGTGTC AGTGAAGAAA TGCCTGTTTC AGAGGTGCCA CCTGGGGGTG ATGAGCGAGG
2881 GCGAAGCCAA AGTCAAACAC TGCGCCTCTA CTGAGACGGG CTGCTTTGTG CTGATCAAGG
2941 GCAATGCCCA AGTCAAGCAT AACATGATCT GTGGGGCCTC GGATGAGCGC GGCTACCAGA
3001 TGCTGACCTG CGCCGGTGGG AACAGCCATA TGCTGGCCAC CGTGCATGTG ACCTCGCACC
3061 CCCGCAAGAC ATGGCCCGAG TTCGAGCACA ACGTCATGAC CCGATGCAAT GTGCACCTGG
3121 GGTCCCGCCG AGGCATGTTT ATGCCCTACC AGTGCAACAT GCAATTTGTG AAGGTGCTGC
3181 TGGAGCCCGA TGCCATGTCC AGAGTGAGCC TGACGGGGGT GTTTGACATG AATGTGGAGC
3241 TGTGGAAAAT TCTGAGATAT GATGAATCCA AGACCAGGTG CCGGGCCTGC GAATGCGGAG
3301 GCAAGCACGC CAGGCTTCAG CCCGTGTGTG TGGAGGTGAC GGAGGACCTG CGACCCGATC
3361 ATTTGGTGTG TTCCTGCAAC GGGACGGAGT TCGGCTCCAG CGGGGAAGAA TCTGACTAGA
3421 GTGAGTAGTG TTTGGGGGAG GTGAGGGGCT TGTATGAGGG GCAGAATGAC TAAAATCTGT
3481 GTTTTCTGT GTGTTGCAGC AGCATGAGCG GAAGCGCCTC CTTTGAGGGA GGGGTATTCA

```

Fig. 8A

SEQ ID No: 4

61/153

```

3541 GCCCTTATCT GACGGGGCGT CTCCCCTCCT GGGCGGGAGT GCGTCAGAAT GTGATGGGAT
3601 CCACGGTGGA CGGCCGGCCC GTGCAGCCCG CGAACTCTTC AACCTTGACC TACGCGACCC
3661 TGAGCTCCTC GTCCGTGGAC GCAGCTGCCG CCGCAGCTGC TGCTTCCGCC GCCAGCGCCG
3721 TGC GCGGAAT GGCCCTGGGC GCCGGCTACT ACAGCTCTCT GGTGGCCAAC TCGACTTCCA
3781 CCAATAATCC CGCCAGCCTG AACGAGGAGA AGCTGCTGCT GCTGATGGCC CAGCTCGAGG
3841 CCCTGACCCA GCGCCTGGGC GAGCTGACCC AGCAGGTGGC TCAGCTGCAG GCGGAGACGC
3901 GGGCCGCGGT TGCCACGGTG AAAACCAAAT AAAAAATGAA TCAATAAATA AACGGAGACG
3961 GTTGTGATT TTAACACAGA GTCTTGAATC TTTATTTGAT TTTTCGCGCG CGGTAGGCCC
4021 TGGACCACCG GTCTCGATCA TTGAGCACCC GGTGGATTTT TTCCAGGACC CGGTAGAGGT
4081 GGGCTTGAT GTTGAGGTAC ATGGGCTATGA GCCCGTCCCG GGGGTGGAGG TAGCTCCATT
4141 GCAGGGCCCTC GTGCTCGGGG GTGGTGTGTT AAATCACCCA GTCATAGCAG GGGCGCAGGG
4201 CGTGGTGCTG CACGATGTCC TTGAGGAGGA GACTGATGGC CACGGGCAGC CCCTTGGTGT
4261 AGGTGTTGAC GAACCTGTTG AGCTGGGAGG GATGCATGCG GGGGAGATG AGATGCATCT
4321 TGGCTGGAT CTTGAGATTG GCGATGTTCC CGCCAGATC CCGCCGGGGG TTCATGTTGT
4381 GCAGGACCAC CAGCACGGTG TATCCGGTGC ACTTGGGAA TTTGTCATGC AACTTGGAAG
4441 GGAAGGCGTG AAAGAATTG GAGACGCCCT TGTGACCGCC CAGGTTTTCC ATGCATCAT
4501 CCATGATGAT GTTGATGGG CCGTGGGCGG CGGCCTGGGC AAAGACGTTT CGGGGTCCG
4561 ACACATCGTA GTTGTGGTCC TGGGTGAGCT CGTCATAGGC CATTTTAATG AATTTGGGGC
4621 GGAGGGTGCC CGACTGGGGG ACGAAGGTGC CCTCGATCCC GGGGGCGTAG TTGCCCTCGC
4681 AGATCTGCAT CTCCAGGCC TTGAGCTCGG AGGGGGGGAT CATGTCCACC TGCGGGGCGA
4741 TGAAAAAAC GGTTCGCGG GCGGGGGAGA TGAGCTGGGC CGAAAGCAGG TTCCGGAGCA
4801 GCTGGGACTT GCCGAGCCG GTGGGGCCGT AGATGACCCC GATGACCGGC TGCAGGTGTT
4861 AGTTGAGGGA GAGACAGCTG CCGTCTCTCG GGAGGAGGGG GGCCACCTCG TTCATCATCT
4921 CGCGCACATG CATGTTCTCG CGCAGAGTT CCGCCAGGAG GCGCTCGCCC CCCAGCGAGA
4981 GGAGCTCTTG CAGCGAGGCG AAGTTTTTCA CGCGCTGAG CCCGTCGGCC ATGGGCATT
5041 TGGAGAGGGT CTGTTGCAAG AGTTCCAGAC GGTCCAGAG CTCGGTGATG TGCTCTAGGG
5101 CATCTCGATC CAGCAGACCT CCTCGTTTCG CCGGTGGGG CGACTGCGGG AGTAGGCAC
5161 CAGGCGATGG GCGTCCAGCG AGGCCAGGGT CCGGTCTTC CAGGGTCGCA GGGTCCGCGT
5221 CAGCGTGGTC TCCGTCACGG TGAAGGGGTG CCGCCCGGGC TGGGCGCTTG CGAGGGTGCG
5281 CTTCAGGCTC ATCCGGCTGG TCGAGAACCG CTCCCGGTCT GCGCCCTGCG CGTCGCGCAG
5341 GTAGCAATTG AGCATGAGTT CGTAGTTGAG CGCCTCGGCC GCGTGGCCCT TGGCGCGGAG
5401 CTTACCTTTG GAAGTGTGTC CGCAGACGGG ACAGAGGAGG GACTTGAGGG CGTAGAGCTT
5461 GGGGGCGAGG AAGACGGACT CGGGGGCGTA GGGCTCCGCG CCGCAGCTGG CGCAGCGGT
5521 CTCGCACTCC ACGAGCCAGG TGAGGTCGGG CCGGTGGGG TCAAAAACGA GGTTCCTCC
5581 GTGCTTTTTG ATGCGTTTCT TACCTCTGGT CTCCATGAGC TCGTGTCCCC GCTGGGTGAC
5641 AAAGAGGCTG TCCGTGTCCC CGTAGACCGA CTTTATGGGC CGGTCTCGA GCGGGTGCC
5701 GCGGTCTCTG TCGTAGAGGA ACCCCGCCCA CTCCGAGACG AAGGCCCGGG TCCAGGCCAG
5761 CACGAAGGAG GCCACGTGGG AGGGGTACCG GTCGTTGTCC ACCAGCGGGT CCACCTTCTC
5821 CAGGGTATGC AAGCACATGT CCCCCTCGTC CACATCCAGG AAGGTGATTG GCTTGTAAGT
5881 GTAGGCCACG TGACCGGGGG TCCCGGCCGG GGGGGTATAA AAGGGGGCGG GCGCTCTCTC
5941 GTCCTCACTG TCTTCCGGAT CGCTGTCCAG GAGCGCCAGC TGTGGGGTA GGTATTCCCT
6001 CTCGAAGGCT GGCATAACCT CGGCACTCAG GTTGTCAATT TCTAGAAACG AGGAGGATTT
6061 GATATTGACG GTGCCGTTGG AGACGCCTTT CATGAGCCCC TCGTCCATCT GGTGAGAAA
6121 GACGATCTTT TTGTTGTGCA GCTTGGTGGC GAAGGAGCCG TAGAGGGCGT TGGAGAGGAG
6181 CTTGGCGATG GAGCGCATGG TCTGGTTCTT TTCCTTGTG GCGCGCTCCT TGGCGCGGAT
6241 GTTGAGCTGC ACGTACTCGC GCGCCACGCA CTTCCATTG GGAAGACGG TGGTGAGCTC
6301 GTCGGGCACG ATTCTGACCC GCCAGCCGCG GTTGTGCAGG GTGATGAGGT CCACGTTGGT
6361 GGCCACCTCG CCGCGCAGGG GCTCGTTGGT CCAGCAGAGG CGCCCGCCCT TCGCGGAGCA
6421 GAAGGGGGGG AGCGGGTCCA GCATGAGCTC GTCGGGGGGG TCGGCGTCCA CGGTGAAGAT
6481 GCCGGGCGA AGCTCGGGGT CGAAGTAGCT GATGCAGGTG TCCAGATCGT CCAGCGCCGC
6541 TTGCCAGTCG CGCACGGCCA GCGCGCGCTG GTAGGGGCTG AGGGCGGTG CCCAGGGCAT
6601 GGGGTGCGTG AGCGCGGAGG CGTACATGCC GCAGATGTCG TAGACGTAGA GGGGCTCCTC
6661 GAGGACGCGG ATGTAGGTGG GGTAGACGCG CCCCCTCGG ATGCTGGCGC GCACGTAGTC
6721 GTACAGCTCG TGCGAGGGCG CGAGGAGCCC CGTGCCGAGG TGGAGAGTGG TGGGCTCTG
6781 GCGCGGCTAG ACGATCTGGC GGAAGATGGC GTGGGAGTTG GAGGAGATGG TGGGCTCTG
6841 GAAGATGTTG AAGTGGGCGT GGGGCGAGCC GACCGAGTCC CTGATGAAGT GGGCGTAGGA
6901 GTCCTGCAGC TTGGCGACGA GCTCGGCGGT GACGAGGACG TCCAGGGCGC AGTAGTCGAG
6961 GGTCTCTTGG ATGATGTCGT ACTTGAGCTG GCCCTTCTGC TTCCACAGCT CGCGGTTGAG
7021 AAGGAACTCT TCGCGGTCCT TCCAGTACTC TTCGAGGGGG AACCCGTCCT GATCGGCACG
7081 GTAAGAGCCC ACCATGTAGA ACTGGTTGAC GGCCTTGATG GCGCAGCAGC CCTTCTCCAC

```

Fig. 8B

SEQ ID No: 4

62/153

```

7141 GGGGAGGGCG TAAGCTTGTG CGGCCCTTGC GAGGGAGGGT TGGGTGAGGG CGAAGGTGTC
7201 GCGCACCATG ACCTTGAGGA ACTGGTGCTT GAAGTCGAGG TCGTCGCAGC CGCCCTGCTC
7261 CCAGAGCTGG AAGTCCGTGC GCTTCTTGTA GGCGGGGTTG GGCAAAGCGA AAGTAACATC
7321 GTTGAAGAGG ATCTTGCCCG CGCGGGGCAT GAAGTTGCGA GTGATGCGGA AAGGCTGGGG
7381 CACCTCGGCC CGGTTGTTGA TGACCTGGGC GGCGAGGACG ATCTCGTCGA AGCCGTTGAT
7441 GTTGTGCCCC ACATGTAGA GTTCCACGAA TCGCGGGCGG CCCTTAACGT GGGGCAGCTT
7501 CTTGAGCTCG TCGTAGGTGA GCTCGGCGGG GTCGCTGAGC CCGTGCTGCT CGAGGGCCCA
7561 GTCGGCGACG TGGGGGTTGG CGCTGAGGAA GGAAGTCCAG AGATCCACGG CCAGGGCGGT
7621 CTGCAAGCGG TCCCAGTACT GACGGAAC TGCGCCACG GCCATTTTTT CGGGGGTGAC
7681 GCAGTAGAAG GTGCGGGGGT CGCCGTGCCA CCGGTCCAC TTGAGCTGGA GGGCGAGGTC
7741 GTGGGCGAGC TCGACGAGCG GCGGGTCCCC GGAGAGTTTC ATGACCAGCA TGAAGGGGAC
7801 GAGCTGCTTG CCGAAGGACC CCATCCAGGT GTAGGTTTCC ACATCGTAGG TGAGGAAGAG
7861 CCTTCGGTG CGAGGATGCG AGCCGATGGG GAAGAACTGG ATCTCCTGCC ACCAGTTGGA
7921 GGAATGGCTG TTGATGTGAT GGAAGTAGAA ATGCCGACGG CGCGCCGAGC ACTCGTGCTT
7981 GTGTTTATAC AAGCGTCCGC AGTGCTCGCA ACCTGTCAG GGATGCACGT GCTGCACGAG
8041 CTGTACCTGG GTTCCCTTGA CGAGGAATTT CAGTGGGCAG TGGAGCGCTG GCGGCTGCAT
8101 CTGGTGCTGT ACTACGTCTT GGCCATCGCG GTGGCCATCG TCTGCCTCGA TGGTGGTCAT
8161 GCTGACGAGC CCGCGCGGGA GGCAGGTCCA GACTTCGGCT CGGACGGGTC GGAGACGGAG
8221 GACGAGGGCG CGCAGGCCGG AGCTGTCCAG GGTCTGAGA CGCTGCGGAG TCAGGTCAGT
8281 GGGCAGCGGC GGCGCGCGGT TGACTTGACG GAGCTTTTCC AGGGCGCGCG GGAGGTCCAG
8341 ATGGTACTTG ATCTCCACGG CGCCGTTGGT GGCGACGTCC ACGGCTTGCA GGGTCCCGTG
8401 CCCCTGGGCG GCCACCACGG TGCCCCGTTT CTTCTTGGGC GCTGCTTCCA TGCCGGTCAG
8461 AAGCGGCGGC GAGGACGCGC GCGGCGCGG AGGGCGGGCT CGGGACCCGG AGGCAGGGGC
8521 GGCAGGGGCA CGTCGGCGCC GCGCGCGGG AGGTTCTGGT ACTGCGCCCG GAGAAGACTG
8581 GCGTGAGCGA CGACGCGACG GTTGACGTCC TGGATCTGAC GCCTCTGGGT GAAGGCCACG
8641 GGACCCGTGA GTTTGAACCT GAAAGAGAGT TCGACAGAAT CAATCTCGGT ATCGTTGACG
8701 GCGGCCTGCC GCAGGATCTC TTGCACGTCC CCCGAGTTGT CCTGGTAGGC GATCTCGGTC
8761 ATGAAC TGCT CGATCTCCTC CTCCTGAAGG TCTCCGCGGC CGGCGCGCTC GACGGTGGCC
8821 GCGAGGTCGT TGGAGATGCG GCCCATGAGC TGCGAGAAGG CGTTCATGCC GGCCTCGTTC
8881 CAGACGCGGC TGTAGACCAC GGCTCCGTCG GGGTCGCGCG CGCGCATGAC CACCTGGGCG
8941 AGGTTGAGCT CGACGTGGCG CGTGAAGACC CCGTAGTTGC AGAGGCGCTG GTAGAGGTAG
9001 TTGAGCGTGG TGGCGATGTG CTCGGTGACG AAGAAGTACA TGATCCAGCG GCGGAGGGGC
9061 ATCTCGCTGA CGTCGCCAG GGCTTCCAAG CGCTCCATGG CCTCGTAGAA GTCCACGGCG
9121 AAGTTGAAAA ACTGGGAGTT GCGCGCCGAG ACGGTCAACT CCTCCTCCAG AAGACGGATG
9181 AGCTCAGCGA TGGTGGCGCG CACCTCGCGC TCGAAGGCCC CGGGGGGCTC CTCTTCTTCC
9241 ATCTCTTCC TCTCCACTAA CATCTCTTCT ACTTCTCTCT CAGGAGGCGG CGGCGGGGGA
9301 GGGGCCCTGC GTCGCGGGCG GCGCAGGTCG AGACGGTCGA TGAAGCGCTC GATGGTCTCC
9361 CCGCGCCGCG GACGCATGGT CTCGGTGACG GCGCGCCCGT CCTCGCGGGG CCGCAGCGTG
9421 AAGACGCCGC CGCGCATCTC CAGGTGGCCG CCGGGGGGGT CTCCGTTGGG CAGGGAGAGG
9481 GCGCTGACGA TGCATCTTAT CAATTGGCCC GTAGGGACTC CGCGCAAGGA CCTGAGCGTC
9541 TCGAGATCCA CGGGATCCGA AAACCGCTGA ACGAAGGCTT CGAGCCAGTC GCAGTCGCAA
9601 GGTAGGCTGA GCGCGGTTTC TTGTTCTTCG GGGATTTCGG GAGGCGGGCG GCGGATGCTG
9661 CTGGTGATGA AGTTGAAGTA GCGGTCCTG AGACGGCGGA TGGTGGCGAG GAGCACCAGG
9721 TCCTTGGGCC CGGCTTGCTG GATGCGCAGA CGGTGCGCCA TGCCCCAGGC GTGGTCTTGA
9781 CACCTGGCGA GGTCTTGTGA GTAGTCTTGC ATGAGCCGCT CCACGGGCAC CTCCTCTCG
9841 CCCGCGCGGC CGTGATGCG CGTGAGCCCG AACCCGCGCT GGGGCTGGAC GAGCCGAGG
9901 TCGGCGACGA CGCGCTCGGC GAGGATGGCC TGCTGTATCT GGGTGAGGGT GGTCTGGAAG
9961 TCGTCAAGT CGACGAAGCG GTGGTAGGCT CCGGTGTTGA TGGTATAGGA GCAGTTGGCC
10021 ATGACGGACC AGTTGACGGT CTGGTGGCCG GGTGCGACGA GCTCGTGGTA CTTGAGGCGC
10081 GAGTAGGCGC GCGTGTCAA GATGTAGTCG TTGCAGGTGC GCACGAGGTA CTGGTATCCG
10141 ACGAGGAAGT GCGGCGGGCG CTGGCGGTAG AGCGGCCATC GCTCGGTGGC GGGGGCGCCG
10201 GCGCGAGGCT CTTGAGCAT GAGGCGGTGG TAGCCGTAGA GTTACCTGGA CATCCAGGTG
10261 ATGCCGCGCG CGGTGGTGGG GCGCGCGGG AACTCGCGGA CGCGGTTCCA GATGTTGCGC
10321 AGCGGCAGGA AGTAGTTCAT GGTGGCCGCG GTCTGGCCCG TGAGGCGCGC GCAGTCGTGG
10381 ATGCTCTAGA CATACGGGCA AAAACGAAAG CGGTCAGCGG CTCGACTCCG TGGCCTGGAG
10441 CCGTAGCGAA CGGGTTGGGC TGCGCGGTGA CCGCGGTTTC AATCTCGAAT CAGGCTGGAG
10501 CCGCAGCTAA CGTGGTACTG GCACTCCCGT CTCGACCCAA GCCTGCTAAC GAAACCTCCA
10561 GGATACGGAG GCGGGTCGTT TTTTGGCCTT GGTGCTGGT CATGAAAAAC TAGTAAGCGC
10621 GGAAAGCGAC CGCCCGCGAT GGCTCGCTGC CGTAGTCTGG AGAAAGAAAT GCCAGGGTTG
10681 CGTTGCGGTG TGCCCCGGTT CGAGCCTCAG CGCTCGGCGC CGGCCGGATT CCGCGGCTAA

```

Fig. 8C

SEQ ID No: 4

63/153

```

10741 CGTGGGCGTG GCTGCCCCGT CGTTTCCAAG ACCCCTTAGC CAGCCGACTT CTCCAGTTAC
10801 GGAGCGAGCC CCTCTTTTTC TTGTGTTTTT GCCAGATGCA TCCCGTACTG CCGCAGATGC
10861 GCCCCACCC TCCACCTCAA CCGCCCTTAC CGCCGCAGCA GCAGCAACAG CCGGCGCTTC
10921 TGCCCCCGCC CCAGCAGCAG CCAGCCACTA CCGCGGCGGC CGCCGTGAGC GGAGCCGCGC
10981 TTCAGTATGA CCTGGCCTTG GAAGAGGGCG AGGGGCTGGC GCGGCTGGGG GCGTCGTCGC
11041 CGGAGCGGCA CCCGCGCGTG CAGATGAAAA GGGACGCTCG CGAGGCCTAC GTGCCCCAAGC
11101 AGAACCTGTT CAGAGACAGG AGCGGCGAGG AGCCCCGAGGA GATGCGCGCC TCCCGCTTCC
11161 ACGCGGGGCG GGAGCTGCGG CGCGGCTTGG ACCGAAAGCG GGTGCTGAGG GACGAGGATT
11221 TCGAGGCGGA CGAGCTGACG GGGATCAGCC CCGCGCGCGC GCACGTGGCC GCGGCCAACC
11281 TGGTCACGGC GTACGAGCAG ACCGTGAAGC AGGAGAGCAA CTTCCAAAAA TCCTTCAACA
11341 ACCACGTGCG CACGCTGATC GCGCGCGAGG AGGTGACCCT GGGCCTGATG CACCTGCGGG
11401 ACCTGCTGGA GGCCATCGTG CAGAACCCCA CGAGCAAGCC GCTGACGGCG CAGCTGTTTC
11461 TGGTGGTGCA GCACAGTCGG GACAACGAGA CGTTCAGGGA GCGGCTGCTG AATATCACCG
11521 AGCCCCGAGG CCGCTGGCTC CTGGACCTGG TGAACATTCT GCAGAGCATC GTGGTGCAGG
11581 AGCGCGGGCT GCCGCTGTCC GAGAAGCTGG CGGCTATCAA CTTCTCGGTG CTGAGCCTGG
11641 GCAAGTACTA CGCTAGGAAG ATCTACAAGA CCCCCTACGT GCCCATAGAC AAGGAGGTGA
11701 AGATCGACGG GTTTTACATG CGCATGACCC TGAAAGTGCT GACCCTGAGC GACGATCTGG
11761 GGGTGTACCG CAACGACAGG ATGCACCGCG CCGTGAGCGC CAGCCGCCGG CCGCGCTGTA
11821 GCGACCAGGA GCTGATGCAC AGCCTGCAGC GGGCCCTGAC CCGGGCCGGG ACCGAGGGGG
11881 AGAGCTACTT TGACATGGGC GCGGACCTGC GCTGGCAGCC CAGCCGCCGG GCCTTGGAAG
11941 CTGCCGCGCG TTCCCCCTAC GTGGAGGAGG TGGACGATGA GGAGGAGGAG GCGGAGTACC
12001 TGGAAGACTG ATGGCGCGAC CGTATTTTTG CTAGATGCAG CAACAGCCAC CGCCTCCTGA
12061 TCCCGCGATG CGGGCGGCGC TGCAGAGCCA GCCGTCCGGC ATTAACTCCT CGGACGATTG
12121 GACCCAGGCC ATGCAACGCA TCATGCGCCT GACGACCCGC AATCCCGAAG CCTTTAGACA
12181 GCAGCCTCAG GCCAACCGGC TCTCGGCCAT CCTGGAGGCC GTGGTGCCCT CGCGCTCGAA
12241 CCCCACGCAC GAGAAGGTGC TGGCCATCGT GAACGCGCTG GTGGAGAACA AGGCCATCCG
12301 CGGCGACGAG GCCCGGCTGG TGTACAACGC GCTGCTGGAG CGCGTGGCCC GCTACAACAG
12361 CACCAACGTG CAGACGAACC TGGACCGCAT GGTGACCGAC GTGCGCGAGG CCGGTGTCGCA
12421 GCGCGAGCGG TTCCACCGCG AGTCGAACCT GGGCTCCATG GTGGCGCTGA ACGCCTTCCT
12481 GAGCACGCGC CCGCCAACG TGCCCCGGGG CCAGGAGGAC TACACCAACT TCATCAGCGC
12541 GCTGCGGCTG ATGGTGGCCG AGGTGCCCCA GAGCGAGGTG TACCAGTCGG GCGCGGACTA
12601 CTTCTTCCAG ACCAGTCGCC AGGGCTTGCA GACCGTGAAC CTGAGCCAGG CTTTCAAGAA
12661 CTTGCAGGGA CTGTGGGGCG TGCAGGCCCC GGTGCGGGAC CCGCGCACGG TGTGAGCCT
12721 GCTGACCCG AACTCGCGCC TGCTGCTGCT GCTGGTGGCG CCCTTCACGG ACAGCGGCAG
12781 CGTGAGCCGC GACTCGTACC TGGGCTACCT GCTTAACCTG TACCGCGAGG CCATCGGGCA
12841 GGCGCACGTG GACGAGCAGA CCTACCAGGA GATCACCCAC GTGAGCCGCG CGCTGGGCCA
12901 GGAGGACCCG GGCAACCTGG AGGCCACCTT GAACTTCCTG CTGACCAACC GGTGCGAGAA
12961 GATCCCGCCC CAGTACGCGC TGAGCACCGA GGAGGAGCGC ATCCTGCGCT ACGTGACGCA
13021 GAGCGTGGGG CTGTTCTCTGA TGCAGGAGGG GSCCAGCCCC AGCGCCGCGC TCGACATGAC
13081 CGCGCGCAAC ATGGAGCCCA GCATGTACGC TCGCAACCGC CCGTTTCATCA ATAAGCTGAT
13141 GGACTACTTG CATCGGGCGG CCGCCATGAA CTCGGACTAC TTTACCAACG CCATCTTGAA
13201 CCCGCACTGG CTCCCGCCGC CCGGGTTCTA CACGGGCGAG TACGACATGC CCGACCCCAA
13261 CGACGGGTTC CTGTGGGACG ACGTGGACAG CAGCGTGTTC TCGCCGCGCC CCGCCACCAC
13321 CGTGTGGAAG AAAGAGGGCG GGGACCGGCG GCCGTCTCG GCGCTGTCCG GTCGCGCGGG
13381 TGCTGCCGCG GCGGTGCCTG AGGCCGCGAG CCCCTTCCCG AGCCTGCCCT TTTCCGTGAA
13441 CAGCGTGCGC AGCAGCGAGC TGGTTCGGCT GACGCGGCCG CGCCTGCTGG GCGAGGAGGA
13501 GTACCTGAAC GACTCCTTGT TGAGGCCCGA GCGCGAGAAG AACTTCCCCA ATAACGGGAT
13561 AGAGAGCCTG GTGGACAAGA TGAGCCGCTG GAAGACGTAC GCGCACGAGC ACAGGGACGA
13621 GCCCCGAGCT AGCAGCAGCG CAGGCACCCG TAGACGCCAG CGACACGACA GGCAGCGGGG
13681 TCTGGTGTGG GACGATGAGG ATTCCGCCGA CGACAGCAGC GTGTTGGACT TGGGTGGGAG
13741 TGGTGGTGGT AACCCGTTCT CTACTTGCG CCCCCGTATC GGGCGCCTGA TGTAAGAATC
13801 TGAAAAATA AAAAACGGTA CTCACCAAGG CCATGGCGAC CAGCGTCCGT TCTTCTCTGT
13861 TGTTTGTAGT AGTATGATGA GGCGCGTGTA CCCGGAGGGT CCTCTCCCTT CGTACGAGAG
13921 CGTGATGCAG CAGGCGGTGG CCGCGGCGAT GCAGCCCCCG CTGGAGGCGC CTTACGTGCC
13981 CCCGCGGTAC CTGGCGCCTA CGGAGGGGCG GAACAGCATT CGTTACTCGG AGCTGGCACC
14041 CTTGTACGAT ACCACCCGGT TGTACCTGGT GGACAACAAG TCGGCGGACA TCGCCTCGCT
14101 GAACTACCAG AACGACCACA GCAACTTCCT GACCACCGTG GTGCAGAAAC ACGATTTCAC
14161 CCCCACGGAG GCCAGCACCC AGAACATCAA CTTTGACGAG CGCTCGCGGT GGGGCGGCCA
14221 GCTGAAAACC ATCATGCACA CCAACGTGAC GAGTTCATGT ACAGCAACAA
14281 GTTCAAGGCG CGGGTGATGG TCTCGCGCAA GACCCCAAT GGGGTGCGCG TGGATGAGAA

```

Fig. 8D

SEQ ID No: 4

64/153

14341 TTATGATGGT AGTCAGGACG AGCTGACTTA CGAGTGGGTG GAGTTTGAGC TGCCCCGAGGG
 14401 CAACTTCTCG GTGACCATGA CCATCGATCT GATGAACAAC GCCATCATCG ACAACTACTT
 14461 GCGGGTGGGG CGTCAGAACG GGGTGTCTGA GAGCGACATC GCGGTGAAGT TCGACACGCG
 14521 CAACTTCCGG CTGGGCTGGG ACCCGGTGAC CGAGCTGGTG ATGCCGGGCG TGTACACCAA
 14581 CGAGGCCTTC CACCCCGACA TCGTCTGCT GCGCGGCTGC GCGGTGGACT TCACCGAGAG
 14641 CCGCCTCAGC AACCTGCTGG GCATCCGCAA GCGGCAGCCC TTCCAGGAGG GCTTCCAGAT
 14701 CCGTACGAG CACCTGGAGG GGGGCAACAT CCGCGGCTC TTGGATGTCG AAGCCTATGA
 14761 GAAAAGCAAG GAGGAGGCCG CCGCAGCGGC GACCGCAGCC GTGGCCACCG CCTCTACCGA
 14821 GGTGCGGGGC GATAATTTTG CTAGCGCCGC GGCAGTGGCC GAGGCGGCTG AAACCGAAAG
 14881 TAAGATAGTC ATCCAGCCGG TGGAGAAGGA CAGCAAGGAC AGGAGCTACA ACGTGCTCGC
 14941 GGACAAGAAA AACACCGCCT ACCCGAGCTG GTACCTGGCC TACAACCTACG GCGACCCCGA
 15001 GAAGGGCGTG CGCTCCTGGA CGCTGCTCAC CACCTCGGAC GTCACCTGCG GCGTGGAGCA
 15061 AGTCTACTGG TCGCTGCCCC ACATGATGCA AGACCCGGTC ACCTTCCGCT CCACGCGTCA
 15121 AGTTAGCAAC TACCCGGTGG TGGGCGCCGA GCTCCTGCCC GTCTACTCCA AGAGCTTCTT
 15181 CAACGAGCAG GCGTCTACT CGCAGCAGCT GCGCGCCTTC ACCTCGCTCA CGCACGTCTT
 15241 CAACCGCTTC CCCGAGAACC AGATCCTCGT CCGCCCGCCC GCGCCACCA TTACCACCGT
 15301 CAGTGAAAAC GTTCTGCTC TCACAGATCA CGGGACCCTG CCGCTGCGCA ACGTATCCG
 15361 GGGAGTCCAG CGCGTGACCG TCACTGACGC CAGACGCCGC ACCTGCCCCT ACGTCTACAA
 15421 GGCCCTGGGC GTAGTCGCGC CGCGCTCCT CTCGAGCCGC ACCTTCTAAA AAATGTCCAT
 15481 TCTCATCTCG CCCAGTAATA ACACCGGTTG GGGCCTGCGC GCGCCAGCA AGATGTACGG
 15541 AGGCGCTCGC CAACGCTCCA CGCAACACCC CGTGCGCGTG GCGGGGCACT TCCGCGCTCC
 15601 CTGGGGCGCC CTCAAGGGCC GCGTGCGCTC GCGCACCACC GTCGACGACG TGATCGACCA
 15661 GGTGGTGGCC GACGCGCGCA ACTACAGCCC CGCCGCGCG CCGCCTCCA CCGTGGACGC
 15721 CGTCATCGAC AGCGTGGTGG CCGATGCGCG CCGGTACGCC CGCGCCAAGA GCCGCGGCG
 15781 GCGCATCGCC CGCGCGCACC GGAGCACCCC CGCCATGCGC GCGGCGCGAG CCTTGCTGCG
 15841 CAGGGCCAGG CGCACGGGAC GCAGGGCCAT GCTCAGGGCG GCCAGACGCG CGGCCTCCGG
 15901 CAGCAGCAGC GCGGCGAGGA CCCGCGAGCG CGCGGCCACG GCGGCGGCGG CGGCCATCGC
 15961 CAGCATGTCC CGCCCCGCGC GCGGCAACGT GTACTGGGTG CGCGACGCCG CCACCGGTGT
 16021 GCGCGTGCCC GTGCGCACCC GCCCTCCG CACTTGAAGA TGCTGACTTC GCGATGTTGA
 16081 TGTGTCCCAG CGGCGAGGAG GATGTCCAAAG CGCAAATACA AGGAAGAGAT GCTCCAGGTC
 16141 ATCGCGCCTG AGATCTACGG CCGCGCGGTG AAGGAGGAAA GAAAGCCCCG CAAACTGAAG
 16201 CGGGTCAAAA AGGACAAAAA GGAGGAGGAA GATGTGGACG GACTGGTGGG GTTTGTGCGC
 16261 GAGTTCGCCC CCGGCGGCG CGTGCACTGG CCGGGGCGGA AAGTGAAACC GGTGCTGCGG
 16321 CCGGCGACCA CCGTGGTCTT CACGCCCGGC GAGCGTTCCG GCTCCGCCTC CAAGCGCTCC
 16381 TACGACGAGG TGTACGGGGA CGAGGACATC CTCGAGCAGG CCGTTCGAGC TCTGGGCGAG
 16441 TTTGCTTACG GCAAGCGCAG CCGCCCCGCG CCCTTGAAG AGGAGGCGGT GTCCATCCCG
 16501 CTGGACCACG GCAACCCAC CCGCAGCCTG AAGCCGCTGA CCCTGCAGCA GGTGCTGCCG
 16561 AGCGCGGCGC CGCGCCGGGG CTTCAGCGC GAGGGCGGCG AGGATCTGTA CCCGACCATG
 16621 CAGCTGATGG TGCCCAAGCG CCAGAAGCTG GAGGACGTGC TGGAGCACAT GAAGGTGGAC
 16681 CCGGAGGTGC AGCCCGAGGT CAAGGTGCGG CCCATCAAGC AGGTGGCCCC GGGCCTGGGC
 16741 GTGCAGACCG TGGACATCAA GATCCCCACG GAGCCCATGG AAACGCAGAC CGAGCCCGTG
 16801 AAGCCAGCA CCAGACCAT GGAGGTGCAG ACGGATCCCT GGATGCCGCG GCCGCTTCC
 16861 ACCACTCGCC GAAGACGCAA GTACGGGCGC GCCAGCCTGC TGATGCCCCA CTACGCGCTG
 16921 CATCCTTCCA TCATCCCCAC GCGGGGCTAC CGCGGCACGC GCTTCTACCG CGGCTACACC
 16981 AGCAGCCGCC GCAAGACCAC CACCCGCGCG CGCCGTCGTC GCACCCGCGC CAGCAGCACC
 17041 GCGACTTCCG CCGCCGCCCT GGTGCGGAGA GTGTACCGCA GCGGGCGCGA GCCTCTGACC
 17101 CTGCCGCGCG CGCGCTACCA CCGGAGCATC GCCATTTAAC TCTGCCGTCG CCTCTACTT
 17161 GCAGATATGG CCCTCACATG CCGCCTCCGC GTCCCATTA CCGGCTACCG AGGAAGAAAG
 17221 CCGCGCCGTA GAAGGCTGAC GGGGAACGGG CTGCGTCCGC ATCACCACCG GCGGCGGCGC
 17281 GCCATCAGCA AGCGGTGGG GGGGAGCTTC CTGCCCGCGC TGATCCCCAT CATCGCCGCG
 17341 GCGATCGGGG CGATCCCCCG CATAGCTTCC GTGGCGGTGC AGGCCTCTCA GCGCCACTGA
 17401 GACACAGCTT GGAAAATTTG TAATAAAAAA ATGGACTGAC GCTCCTGGTC CTGTGATGTG
 17461 TGTTTTATGA TGGAAGACAT CAATTTTTCG TCCCTGGCAC CGCGACACGG CACGCGGCGC
 17521 TTTATGGGCA CCTGGAGCGA CATCGGCAAC AGCCAACCTGA ACGGGGGCGC CTTCATTTGG
 17581 AGCAGTCTCT GGAGCGGGCT TAAGAATTTT GGGTCCACGC TCAAAACCTA TGGCAACAAG
 17641 GCGTGGAACA GCAGCACAGG GCAGGCGCTG AGGGAAAAGC TGAAAGAGCA GAACTTCCAG
 17701 CAGAAAGTGG TCGATGGCCT GGCCTCGGGC ATCAACGGGG TGGTGGACCT GGCCAACCAG
 17761 GCGGTGCAGA AACAGATCAA CAGCCGCTG GACGCGGTCC CGCCGCGGG GTCCGTGGAG
 17821 ATGCCCCAGG TGGAGGAGGA GCTGCCCTCC CTGGACAAGC GCGGCGACAA GCGACCGCT
 17881 CCGGACGCGG AGGAGACGCT GCTGACGCAC ACGGACGAGC CGCCCCGTA CGAGGAGGCC

Fig. 8E

SEQ ID No: 4

65/153

```

17941 GTGAAACTGG GTCTGCCCAC CACGCGGCCC GTGGCGCCTC TGGCCACCGG GGTGCTGAAA
18001 CCCAGCAGCA GCAGCCAGCC CGCGACCCTG GACTTGCCCTC CGCCTGCTTC CCGCCCCCTCC
18061 ACAGTGGCTA AGCCCCTGCC GCCGTGGGCC GTCGCGTCGC GCGCCCCCGG AGGCCGCCCC
18121 CAGGCGAACT GGCAGAGCAC TCTGAACAGC ATCGTGGGTC TGGGAGTGCA GAGTGTGAAG
18181 CGCCGCCGCT GCTATTAAAA GACACTGTAG CGCTTAACTT GCTTGTCTGT GTGTATATGT
18241 ATGTCCGCCG ACCAGAAGGA GGAAGAGGCG CGTCGCCGAG TTGCAAGATG GCCACCCCAT
18301 CGATGCTGCC CCAGTGGGCG TACATGCACA TCGCCGGACA GGACGCTTCG GAGTACCTGA
18361 GTCCGGGTCT GGTGCAGTTC GCCCGCGCCA CAGACACCTA CTTCAGTCTG GGAACAAGT
18421 TTAGGAACCC CACGGTGGCG CCCACGCACG ATGTGACCAC CGACCGCAGC CAGCGGCTGA
18481 CGCTGCGCTT CGTGCCCGTG GACCGCGAGG ACAACACCTA CTCGTACAAA GTGCGCTACA
18541 CGCTGGCCGT GGGCGACAAC CGCGTGCTGG ACATGGCCAG CACCTACTTT GACATCCGCG
18601 GCGTGTGGA TCGGGGGCCC AGCTTCAAAC CCTACTCCGG CACCGCCTAC AACAGCCTGG
18661 CTCCCAAGGG AGCGCCCAAC ACTTGCCAGT GGACATATAA AGCTGGTGAT ACTGATACAG
18721 AAAAAACCTA TACATATGGA AATGCACCTG TGCAAGGCAT TAGCATTACA AAGGATGGTA
18781 TTCAACTTGG AACTGACAGC GATGGTCAGG CAATCTATGC AGACGAAACT TATCAACCAG
18841 AGCCTCAAGT GGGTGATGCT GAATGGCATG ACATCACTGG TACTGATGAA AAATATGGAG
18901 GCAGAGCTCT TAAGCCTGAC ACCAAAAATGA AGCCTTGCTA TGGTTCTTTT GCCAAGCCTA
18961 CCAATAAAGA AGGAGGCCAG GCAAAATGTGA AAACCGAAAC AGGCGGTACC AAAGAATATG
19021 ACATTGACAT GGCATTCTTC GATAATCGAA GTGCAGCTGC CGCCGGCCTA GCCCGGAAA
19081 TTGTTTGTA TACTGAGAAT GTGGATCTGG AAATCCAGA TACCATATT GTATACAAGG
19141 CAGGTACAGA TGACAGTAGC TCTTCTATCA ATTTGGGTCA GCAGTCCATG CCAACAGAC
19201 CCAACTACAT TGGCTTCAGA GACAACTTTA TCGGTCTGAT GTACTACAAC AGCACTGGCA
19261 ATATGGGTGT ACTGGCTGGA CAGGCTCCCG AGCTGAATGC TGTGGTGGAC TTGCAGGACA
19321 GAAACACCGA ACTGTCTTAC CAGCTCTTGC TTGACTCTCT GGGTGACAGA ACCAGGTATT
19381 TCAGTATGTG GAATCAGGCG GTGGACAGTT ATGACCCCGA TGTGCGCATT ATTGAAAATC
19441 ACGGTGTGGA GGATGAACCT CTAACCTATT GCTTCCCCCT GGATGCTGTG GGTAGAAGCTG
19501 ATACTTACCA GGAATTAAG GCCAATGGTG ATAATCAAAC CACCTGGACC AAAGATGATA
19561 CTGTTAATGA TGCTAATGAA TTGGGCAAGG GCAATCCTTT CGCCATGGAG ATCAACATCC
19621 AGGCCAACCT GTGGCGGAAC TTCCTCTACG CGAACGTGGC GCTGTACCTG CCCGACTCCT
19681 ACAAGTACAC GCCGGCCAAC ATCAGCGTGC CCACCAACAC CAACACCTAC GATTACATGA
19741 ACGGCCGCGT GGTGGCGCCC TCGTGGTGG AGCCTACAT CAACATCGGG GCGCGCTGGT
19801 CGCTGGACCC CATGGACAAC GTCAACCCCT TCAACCACCA CCGCAACGCG GGCCTGCGAT
19861 ACCGCTCCAT GCTCCTGGGC AACGGGCGCT ACGTGCCCTT CCACATCCAG GTGCCCAAA
19921 AGTTTTTCGC CATCAAGAGC CTCCTGCTCC TGCCCGGGTC CTACACCTAC GAGTGGAACT
19981 TCCGCAAGGA CGTCAACATG ATCTTGAGA GCTCCCTCGG CAACGACCTG CGCACGGACG
20041 GGGCTCCAT CGCCTTCACC AGCATCAACC TCTACGCCAC CTTCCTCCCC ATGGCGCACA
20101 ACACCGCCTC CACGCTCGAG GCCATGCTGC GCAACGACAC CAACGACCAG TCCTTCAACG
20161 ACTACTCTC GCGGCCAAC ATGCTCTACC CCATCCCGGC CAACGCCACC AACGTGCCCA
20221 TCTCCATCCC CTCGCGCAAC TGGCCGCTG TCCGCGGTG GTCCCTCAGC CGCTTCAAG
20281 CCCGCGAGAC GCCCTCGCTC GGCTCCGGGT TCGACCCCTA CTTCGTCTAC TCGGCTTCCA
20341 TCCCTACCT CGACGGCACC TTCTACCTCA ACCACACCTT CAAGAAGGTC TCCATCACCT
20401 TCGACTCCTC CGTCAGCTGG CCCGGCAACG ACCGCTCCTT GACGCCCCAAC GAGTTGAAA
20461 TCAAGCGCAC CGTCGACGGA GAGGGGTACA ACGTGGCCCA GTGCAACATG ACCAAGGACT
20521 GGTTCTTGTT CCAGATGCTG GCCCACTACA ACATCGGCTA CCAGGGCTTC TACGTGCCCCG
20581 AGGGCTACAA GGACCGCATG TACTCCTTCT TCCGCAACTT CCAGCCCATG AGCCGCCAGG
20641 TCGTGGACGA GGTCAACTAC AAGGACTACC AGGCCGTAC CCTGGCCTAC CAGCACAACA
20701 ACTCGGGCTT CGTCGGCTAC CTCGCGCCCA CCATGCGCCA GGGCCAGCCC TACCCGCCCA
20761 ACTACCCCTA CCCGCTCATC GGCAAGAGCG CCGTCGCCAG CGTCACCCAG AAAAAGTTCC
20821 TCTGCGACCG GGTATGTGG CGCATCCCCT TCTCCAGCAA CTTCATGTCC ATGGGCGCGC
20881 TCACCGACCT CGGCCAGAAC ATGCTCTACG CCAACTCCGC CCACGCGCTA GACATGAATT
20941 TCGAAGTCGA CCCCATGGAT GAGTCCACCT TTCTCTATGT TGTCTTCGAA GTCTTCGACG
21001 TCGTCCGAGT GCACGACCC CACCGCGGCG TCATCGAGGC CGTCTACCTG CGCACGCCCT
21061 TCTCGGCCGG CAACGCCACC ACCTAAGCCT CTTGCTTCTT GCAAGATGAC GGCCTGCGCG
21121 GGCTCCGGCG AGCAGGAGCT CAGGGCCATC CTCCGCGACC TGGGCTGCGG CCCCTGCTTC
21181 CTGGGCACCT TCGACAAGCG CTTCCCGGGA TTCATGGCCC CGCACAAGCT GGCCTGCGCC
21241 ATCGTCAACA CCGCCGGCCG CGAGACCGGG GGCGAGCACT GGCTGGCCTT CGCCTGGAAC
21301 CCGCGCTCCC ACACCTGCTA CCTCTTCGAC CCCTTCGGGT TCTCGGACGA GCGCCTCAAG
21361 CAGATCTACC AGTTCGAGTA CGAGGGCCTG CTGCGTCGCA GCGCCCTGGC CACCGAGGAC
21421 CGCTGCGTCA CCCTGGAAA GTCCACCCAG ACCGTGCAGG GTCCGCGCTC GGCCGCTGCG
21481 GGGCTCTTCT GCTGCATGTT CCTGCACGCC TTCGTGCACT GGCCCCACCG CCCCATGGAC

```

Fig. 8F

SEQ ID No: 4

66/153

```

21541 AAGAACCCCA CCATGAACCTT GCTGACGGGG GTGCCCAACG GCATGCTCCA GTGCCCCCAG
21601 GTGGAACCCCA CCTTGCGCCG CAACCAGGAG GCGCTCTACC GCTTCTCAA CGCCCACTCC
21661 GCCTACTTTC GCTCCACCG CGCGCGCATC GAGAAGGCCA CCGCTTCGA CCGCATGAAT
21721 CAAGACATGT AATCCGGTGT GTGTATGTGA ATGCTTTATT CATCATAATA AACAGCACAT
21781 GTTTATGCCA CCTTCTCTGA GGCTCTGACT TTATTTAGAA ATCGAAGGGG TTCTGCCGGC
21841 TCTCGGCATG GCCCCGCGGC AGGGATACGT TGCAGAACTG GTACTTGGGC AGCCACTTGA
21901 ACTCGGGGAT CAGCAGCTTC GGCACGGGGA GGTGCGGGAA CGAGTCGCTC CACAGCTTGC
21961 GCGTGAGTTG CAGGGCGCCC AGCAGGTCGG GCGCGGAGAT CTTGAAATCG CAGTTGGGAC
22021 CCGCGTTCTG CGCGCGAGAG TTACGGTACA CCGGGTTGCA GCACTGGAAC ACCATCAGGG
22081 CCGGGTGCTT CACGCTCGCC AGCACCCTCG CGTCGGTGAT GCCCTCCACG TCCAGATCCT
22141 CCGCGTTGGC CATCCCGAAG GGGGTCATCT TGCAGGTCTG CCGCCCCATG CTGGGCACGC
22201 AGCCGGGCTT GTGGTTGCAA TCGCAGTGCA GGGGGATCAG CATCATCTGG GCCTGCTCGG
22261 AGCTCATGCC CGGGTACATG GCCTTCATGA AAGCCTCCAG CTGGCGGAAG GCCTGCTGCG
22321 CCTTGCCGCC CTCGGTGAAG AAGACCCCGC AGGACTTGCT AGAGAACTGG TTGGTGGCGC
22381 AGCCAGCGTC GTGCACGCAG CAGCGCGCGT CGTTGTTGGC CAGCTGCACC ACGTGCGCC
22441 CCCAGCGGTT CTGGGTGATC TTGGCCCGGT CCGGGTTCTC CTTACGCGCG CGCTGCCCGT
22501 TCTCGCTCGC CACATCCATC TCGATCGTGT GCTCCTCTG GATCATCACG GTCCCGTGCA
22561 GGCACCGCAG CTTGCCCTCG GCCTCGGTGC ACCCGTGCAG CCACAGCGCG CAGCCGGTGC
22621 TCTCCAGTT CTTGTGGGCG ATCTGGGAGT GCGAGTGCAC GAAGCCCTGC AGGAAGCGGC
22681 CCATCATCGT GGTACGGGTC TTGTGCTGG TGAAGGTCAG CGGAATGCCG CCGTGCTCCT
22741 CGTTCACATA CAGGTGGCAG ATACGGCGGT ACACCTCGCC CTGCTCGGGC ATCAGCTGGA
22801 AGCGGACTT CAGTCTGCTC TCCACGCGGT ACCGGTCCAT CAGCAGCGTC ATCACTTCCA
22861 TGCCCTTCTC CCAGGCCGAA ACGATCGGCA GGCTCAGGGG GTTCTTCACC GTTGTCTCT
22921 TAGTCGCCGC CGCCGAAGTC AGGGGGTCGT TCTCGTCCAG GGTCTCAAAC ACTCGCTTGC
22981 CGTCTTCTC GGTGATGCGC ACGGGGGGAA AGCTGAAGCC CACGGCCGCC AGCTCCTCCT
23041 CGGCCTGCCT TTCGTCTCG CTGTCTTGGC TGATGTCTTG CAAAGGCACA TGCTTGGTCT
23101 TGCGGGGTTT CTTTTTGGGC GGCAGAGGCG GCGGCGGAGA CGTGCTGGGC GAGCGCGAGT
23161 TCTCGCTCAC CACGACTATT TCTTCTCTTT GGCCGTCGTC CGAGACCACG CCGCGGTAGG
23221 CATGCCCTTT CTGGGCGAGA GCGGAGGCG ACGGGCTCTC GCGGTTCGGC GGGCGGCTGG
23281 CAGAGCCCTT TCCGCGTTTC GGGGTGCGT CCTGGCGGCG CTGCTCTGAC TGACTTCTCT
23341 CGCGGCCGCG CATTGTGTTT TCCTAGGAGT CAAGCATGGA GACTCAGCCA TCGTCGCCAA
23401 CATCGCCATC TGCCCCCGCC GCCGCCGACG AGAACCAGCA GCAGCAGAAT GAAAGCTTAA
23461 CCGCCCCGCC GCCCAGCCCC ACCTCCGACG CCGCAGCCCC AGACATGCAA GAGATGGAGG
23521 AATCCATCGA GATTGACCTG GGCTACGTGA CGCCCGCGGA GCACGAGGAG GAGCTGGCAG
23581 CGCGCTTTC AGCCCCGGA GAGAACCACC AAGAGCAGCC AGAGCAGGAA GCAGAGAGCG
23641 AGCAGAACCA GGCTGGGCTC GAGCATGGCG ACTACCTGAG CCGGGCAGAG GACGTGCTCA
23701 TCAAGCATCT GGCCCCGCAA TGCATCATCG TCAAGGACGC GCTGCTCGAC CGCGCCGAGG
23761 TGCCCTCAG CGTGGCGGAG CTCAGCCGCG CCTACGAGCG CAACCTCTTC TCGCCGCGCG
23821 TGCCCCCAA GCGCCAGCCC AACGGCACCT GCGAGCCCAA CCGCGCCTC AACTTCTACC
23881 CGGTCTTCGC GGTGCCCGAG GCCCTGGCCA CCTACCACCT CTTTTTCAAG AACCAGGGA
23941 TCCCCGTCTC CTGCCGCGCC AACCGCACCC GCGCCGACGC CCTGCTCAAC CTGGGCCCCG
24001 GCGCCCGCCT ACCTGATATC GCCTCCTTGG AAGAGGTTCC CAAGATCTTC GAGGGTCTGG
24061 GCAGCGACGA GACTCGGGCC GCGAACGCTC TGCAAGGAAG CGGAGAGGAG CATGAGCACC
24121 ACAGCGCCCT GGTGGAGTTG GAAGGCGACA ACGCGCGCCT GGCGGTCTCT AAGCGCACGG
24181 TCGAGCTGAC CCACTTCGCC TACCCGCGCG TCAACCTGCC CCCCAGGTC ATGAGCGCCG
24241 TCATGGACCA GGTGCTCATC AAGCGCGCCT CGCCCTCTC GGAGGAGGAG ATGCAGGACC
24301 CCGAGAGCTC GGACGAGGCG AAGCCCGTGG TCAGCGACGA GCAGCTGGCG CGCTGGCTGG
24361 GAGCGAGTAG CACCCCCCAG AGCCTGGAAG AGCGGCGCAA GCTCATGATG GCCGTGGTCC
24421 TGGTGACCGT GGAGCTGGAG TGTCTGCGCC GCTTCTTCGC CGACGCGGAG ACCCTGCGCA
24481 AGGTGAGGA GAACCTGCAC TACCTGTTCA GACACGGGTT CGTGCGCCAG GCCTGCAAGA
24541 TCTCCAACGT GGAGCTGACC AACCTGCTCT CCTACATGGG CATCCTGCAC GAGAACCGCC
24601 TGGGGCAGAA CGTGCTGCAC ACCACCCTGC GCGGGGAGGC CCGCCGCGAC TACATCCGCG
24661 ACTGCGTCTA CCTGTACCTC TGCCACACCT GGCAGACGGG CATGGGCGTG TGGCAGCAGT
24721 GCCTGGAGGA GCAGAACCTG AAAGAGCTCT GCAAGCTCCT GCAGAAAGAAC CTCAAGGCC
24781 TGTGGACCGG GTTCGACGAG CGCACCACCG CCGCGGACCT GGCCGACCTC ATCTTCCCCG
24841 AGCGCCTGCG GCTGACGCTG CGCAACGGGC TGCCCGACTT TATGAGCCAA AGCATGTTGC
24901 AAAACTTTTC CTCTTTCATC CTCGAACGCT CCGGGATCCT GCCCGCCACC TGCTCCGCGC
24961 TGCCCTCGGA CTTCTGTGCG CTGACCTTCC GCGAGTGCCC CCGCGCGCTC TGGAGCCACT
25021 GCTACCTGCT GCGCCTGGCC AACTACCTGG CTAACACTC GGACGTGATC GAGGACGTCA
25081 GCGCGAGGGG CCTGCTCGAG TGCCACTGCC GCTGCAACCT CTGCACGCC CACCGCTCCC

```

Fig. 8G

SEQ ID No: 4

67/153

```

25141 TGGCCTGCAA CCCCCAGCTG CTGAGCGAGA CCCAGATCAT CGGCACCTTC GAGTTGCAAG
25201 GCCCCGGCGA GGGCAAGGGG GGTCTGAAAC TCACCCCGGG GCTGTGGACC TCGGCCTACT
25261 TGC GCAAGTT CGTGCCCGAG GACTACCATC CCTTCGAGAT CAGGTTCTAC GAGGACCAAT
25321 CCCAGCCGCC CAAGGCCGAG CTGTGCGCCT GCGTCATCAC CCAGGGGGCC ATCCTGGCCC
25381 AATTGCAAGC CATCCAGAAA TCCCGCCAAG AATTTCTGCT GAAAAAGGGC CACGGGGTCT
25441 ACTTGACCCC CCAGACCGGA GAGGAGCTCA ACCCCAGCTT CCCCAGGAT GCCCCGAGGA
25501 AGCAGCAAGA AGCTGAAAGT GGAGCTGCCG CCGCCGCCGG AGGATTTGGA GGAAGACTGG
25561 GAGAGCAGTC AGGCAGAGGA GGAGGAGATG GAAGACTGGG ACAGCACTCA GGCAGAGGAG
25621 GACAGCCTGC AAGACAGTCT GGAGGAGGAA GACGAGGTGG AGGAGGCAGA GGAAGAAGCA
25681 GCCGCCGCCA GACCGTCGTC CTCGGCGGAG GAGGAGAAAG CAAGCAGCAC GGATACCATC
25741 TCCGCTCCGG GTCGGGGTCG CCGCGGCCGG GCCCACAGTA GATGGGACGA GTCCCTGGCG
25801 TTCCCGAACC CCACCACCCA GACCGGTAAG AAGGAGCGGC AGGGATACAA GTCCCTGGCG
25861 GGGCACA AAAA ACGCCATCGT CTCTGCTTG CAAGCCTGCG GGGGCAACAT CTCTTCACC
25921 CGGCGCTACC TGCTCTTCCA CCGCGGGGTG AACTTCCCC GCAACATCTT GCATTACTAC
25981 CGTCACTCC ACAGCCCCTA CTACTGTTTC CAAGAAGAGG CAGAAACCCA GCAGCAGCAG
26041 CAGCAGCAGA AAACAGCGG CAGCAGCTAG AAAATCCACA GCGGCGGCAG GTGGACTGAG
26101 GATCGCGGCG AACGAGCCGG CGCAGCGGAG GGAGCTGAGG AACC GGATCT TTCCCAACCT
26161 CTATGCCATC TTCCAGCAGA GTCGGGGGCA AGAGCAGGAA CTGAAAGTCA AGAACCGTTC
26221 TCTGCGCTCG CTCACCCGCA GTTGCTGTGA TCACAAGAGC GAAGACCAAC TTACGCGCAC
26281 TCTCGAGGAC GCCGAGGCTC TCTTCAACAA GTACTGCGCG CTCACTCTTA AAGAGTAGCC
26341 CGCGCCCGCC CACACACGGA AAAAGGCGGG AATTACGTCA CCACCTGCGC CTTTCGCCCG
26401 ACCATCATCA TGAGCAAAGA GATTCACACG CCTTACATGT GGAGCTACCA GCCCAGATG
26461 GGCCTGGCCG CCGCGCCGCG CCAGGACTAC TCCACCCGCA TGAAC TGGCT CAGTGCCGGG
26521 CCCGCGATGA TCTCACGGGT GAATGACATC CGCGCCACCC GAAACCAGAT ACTCCTAGAA
26581 CAGTCAGCGA TCACCGCCAC GCCCCGCCAT CACCTTAATC CGCGTAATTG GCCCGCCGCC
26641 CTGGTGTACC AGGAAATTC CCAGCCACG ACCGTACTAC TTCCGCGAGA CGCCAGGCC
26701 GAAGTCCAGC TGACTAACTC AGGTGTCCAG CTGGCCGGCG GCGCCGCCCT GTGTGCTCAC
26761 CGCCCCGCTC AGGGTATAAA GCGGCTGGTG ATCCGAGGCA GAGGCACACA GCTCAACGAC
26821 GAGGTGGTGA GCTCTTCGCT GGGTCTGCGA CCTGACGGAG TCTTCCAAC TCGCCGATCG
26881 GGGAGATCTT CTTTCACGCC TCGTCAGGCC GTCCTGACTT TGGAGAGTTC GTCTCGCAG
26941 CCCC GCTCGG GTGGCATCGG CACTCTCCAG TTCGTGGAGG AGTTCCTCC CTCGGTCTAC
27001 TTCAACCCCT TCTCCGCTC CCCC GCGCAC TACCCGACG AGTTCATCC GAACCTGAC
27061 GCCATCAGCG AGTCGGTGGA CGGCTACGAT TGAATGTCCC ATGGTGGCGC GGCTGACCTA
27121 GCTCGGCTTC GACACCTGGA CCACTGCCGC CGCTTCCGCT GCTTCGCTCG GGATCTCGCC
27181 GAGTTTGCTT ACTTTGAGCT GCCC GAGGAG CACCCTCAGG GCCCGGCCA CGGAGTGCGG
27241 ATCGTCTGCG AAGGGGGTCT CGACTCCCAC CTGCTTCGGA TCTTCAGCCA GCGTCCGATC
27301 CTGGCCGAGC GCGAGCAAGG ACAGACTCCT CTGACCCCTGT ACTGCATCTG CAACCACCCC
27361 GGCTGTCATG AAAGTCTTTG TTGTCTGTG TGTACTGAGT ATAATAAAAG CTGAGATCAG
27421 CGACTACTCC GGACTTCCGT GTGTTCTCTG TATCAACCAG TCCCTGTTCT TCACCGGGA
27481 CGAGACCGAG CTCCAGCTCC AGTGTAAGCC CCACAAGAAG TACCTCACCT GGCTGTTCCA
27541 GGGCTCTCCG ATCGCCGTTG TCAACCACTG CGACAACGAC GGAGTCTTGC TGAGCGGCC
27601 TGCCAACCTT ACTTTTCCA CCCG CAGAAG CAAGTCCAG CTCTTCCAAC CCTTCTCCC
27661 CGGGACCTAT CAGTGCGTCT CGGGACCTTG CCATCACACC TTCCACCTGA TCCCGAATAC
27721 CACAGCGTCG CTCCCCGCTA CTAACAACCA AACTACCCAC CAACGCCACC GTCGCGACCT
27781 TTCTCTGGG TCTAATACCA CTACCGGAGG TGAGCTCCGA GGTGACCAA CCTCTGGGAT
27841 TTACTACGGC CCCTGGGAGG TGGTAGGGTT AATAGCGCTA GGCTAGTTG CGGTGGGCT
27901 TTTGGCTCTC TGCTACCTAT ACCTCCCTTG CTGTTCTGTA TTAGTGGTGC TGTGTTGCTG
27961 GTTTAAGAAA TGGGGAAGAT CACCCTAGTG AGCTGCGGTG TGCTGGTGGC GGTGGTGGCT
28021 TCGATTGTGG GACTGGGCGG CGCGGCTGTA GTGAAGGAGA AGGCCGATCC CTGCTTGCAT
28081 TTCAATCCCG ACAAATGCCA GCTGAGTTT CAGCCCGATG GCAATCGGTG CGCGGTGCTG
28141 ATCAAGTGCG GATGGGAATG CGAGAACGTG AGAATCGAGT ACAATAACAA GACTCGGAAC
28201 AATACTCTCG CGTCCGTGTG GCAGCCCGGG GACCCCGAGT GGTACACCGT CTCGTGCCCC
28261 GGTGCTGACG GCTCCCCGCG CACCGTGAAT AATACTTTCA TTTTTCGCA CATGTGCGAC
28321 ACGGTCATGT GGATGAGCAA GCAGTACGAT ATGTGGCCCC CCACGAAGGA GAACATCGTG
28381 GTCTTCTCCA TCGCTTACAG CGTGTGCACG GCGCTAATCA CCGCTATCGT GTGCCTGAGC
28441 ATTACATGTC TCATCGCTAT TCGCCCCAGA AATAATGCCG AAAAAGAAAA ACAGCCATAA
28501 CACGTTTTTT CACACACCTT TTT CAGACCA TGGCCTCTGT TAAATTTTTG CTTTTATTTG
28561 CCAGTCTCAT TGCCGTCAAT CATGGAATGA GTAATGAGAA AATTACTATT TACACTGGCA
28621 CTAATCACAC ATTGAAAGGT CCAGAAAAAG CCACAGAAGT TTCATGGTAT TGTATTTTTA
28681 ATGAATCAGA TGTATCTACT GAACTCTGTG GAAACAATAA CAAAAAAAT GAGAGCATTA

```

Fig. 8H

SEQ ID No: 4

68/153

```

28741 CTCTCATCAA GTTTCAATGT GGATCTGACT TAACCCTAAT TAACATCACT AGAGACTATG
28801 TAGGTATGTA TTATGGAAC ACAGCAGGCA TTTCGGACAT GGAATTTTAT CAAGTTTCTG
28861 TGTCTGAACC CACCACGCCT AGAATGACCA CAACCACAAA AACTACACCT GTTACCACTA
28921 TACAGCTCAC TACCAATGGC TTTCTTGCCA TGCTTCAAGT GGCTGAAAAT AGCACCAGCA
28981 TTCAACCCAC CCCACCCAGT GAGGAAATTC CCAGATCCAT GATTGGCATT ATTGTTGCTG
29041 TAGTGGTGTG CATGTTGATC ATCGCCTTGT GCATGGTGTA CTATGCCTTC TGCTACAGAA
29101 AGCACAGACT GAACGACAAG CTGGAACACT TACTAAGTGT TGAATTTTAA TTTTGTAGAA
29161 CCATGAAGAT CCTAGGCCTT TTAGTTTTTT CTATCATTAC CTCTGCTCTA TGCAATTCCTG
29221 ACAATGAGGA CGTTACTGTC GTTGTCGGAT CAAATTATAC ACTAAAAGGT CCAGCAAAAG
29281 GTATGCTTTC GTGGTATTGT TGGTTCGGAA CTGACGAGCA ACAGACAGAA CTTTGCAATG
29341 CTCAAAAAGG CAAAACCTCA AATTCTAAAA TCTCTAATTA TCAATGCAAT CAGACTGACT
29401 TAGTATTGCT CAATGTCACG AAAGCATATG CTGGCAGTTA CACCTGCCCT GGAGATGATG
29461 CCGACAATAT GATTTTTTAC AAAGTGGAAG TGGTTGATCC CACTACTCCA CCGCCCACCA
29521 CCACAACCTAC TCATACCACA CACACAGAAC AAACACCAGA GGCAGCAGAA GCAGAGTTGG
29581 CCTTCCAGGT TCACGGAGAT TCCTTTGCTG TCAATACCCC TACACCCGAT CAGCGGTGTC
29641 CGGGGCTGCT CGTCAGCGGC ATTGTCGGTG TGCTTTCGGG ATTAGCAGTC ATAATCATCT
29701 GCATGCTTTC TTTTGCTTGC TGCTATAGAA GGCTTTACCG AAAAAAATCA GACCCACTGC
29761 TGAACCTCTA TGTTTAATTT TTTCCAGAGC CATGAAGGCA GTTAGCGCTC TAGTTTTTTG
29821 TTCTTTGATT GGCATTGTTT TTAGTGCTGG GTTTTTGAAA AATCTTACCA TTTATGAAGG
29881 TGAGAATGCC ACTCTAGTGG GCATCAGTGG TCAAAATGTC AGCTGGCTAA AATACCATCT
29941 AGATGGGTGG AAAGACATTT GCGATTGGAA TGCTACTGTG TATACATGTA ATGGAGTTAA
30001 CCTCACCATT ACTAATGCCA CCCAAGATCA GAATGGTAGG TTTAAGGGCC AGAGTTTCAC
30061 TAGAAATAAT GGGTATGAAT CCCATAACAT GTTATCTAT GACGTCACCTG TCATCAGAAA
30121 TGAGACTGCC ACCACCACAC AGATGCCCCA TACACACAGT TCTACCACTA CTACCATGCA
30181 AACCACACAG ACAACCACTA CATCAACTCA GCATATGACC ACCACTACAG CAGCAAAGCC
30241 AAGTAGTGCA GCGCCTCAGC CCCAGGCTTT GGCTTTGAAA GCTGCACAAC CTAGTACAAC
30301 TACTAGGACC AATGAGCAGA CTACTGAATT TTTGTCCACT GTCGAGAGCC ACACCACAGC
30361 TACCTCCAGT GCCTTCTCTA GCACCGCCAA TCTCTCCTCG CTTTCTCTTA CACCAATCAG
30421 TCCCGCTACT ACTCCCACCC CAGCTCTTCT CCCCCTCCCT CTGAAGCAAA CTGAGGACAG
30481 CGGCATGCAA TGGCAGATCA CCTGCTCAT TGTGATCGGG TTGGTCATCC TGGCCGTGTT
30541 GCTCTACTAC ATCTTCTGCC GCCGACTTCC CAACGCGCAC CGCAAACCGG CCTACAAGCC
30601 CATCGTTATC GGGCAGCCGG AGCCGCTTCA GGTGGAAGGG GGTCTAAGGA ATCTTCTCTT
30661 CTCTTTTACA GTATGGTGAT TGAACATATG TTCTTAGACA ATTCTTGATC ACTATTCTTA
30721 TCTGCCTCCT CCAAGTCTGT GCCACCCTCG CTCTGGTGGC CAACGCCAGT CCAGACTGTA
30781 TTGGGCCCCT CGCCTCCTAC GTGCTCTTTG CCTTCATCAC CTGCATCTGC TGCTGTAGCA
30841 TAGTCTGCCT GCTTATCACC TTCTTCCAGT TCATTGACTG GATCTTTGTG CGCATCGCCT
30901 ACCTGCGCCA CCACCCCAG TACCGCGACC AGCGAGTGGC GCGGCTGCTC AGGCTCCTCT
30961 GATAAGCATG CGGGCTCTGC TACTTCTCGC GCTTCTGCTG TTAGTGCTCC CCCGCCCCGT
31021 CGACCCCCGG TCCCCCACTC AGTCCCCCGA AGAGGTCCGC AAATGCAAAAT TCCAAGAACC
31081 CTGGAAATTC CTCAAATGCT ACCGCCAAAA ATCAGACATG CTTCCCAGCT GGATCATGAT
31141 CATTGGGATC GTGAACATTC TGGCCTGCAC CCTCATCTCC TTTGTGATTT ACCCCTGCTT
31201 TGACTTTGGT TGGAACTCGC CAGAGGCGCT CTATCTCCCG CCTGAACCTG ACACACCACC
31261 ACAGCAACCT CAGGCACACG CACTACCACC ACCACAGCCT AGGCCACAAT ACATGCCCAT
31321 ATTAGACTAT GAGGCCGAGC CACAGCGACC CATGCTCCCC GCTATTAGTT ACTTCAATCT
31381 AACC GGCGGA GATGACTGAC CCACTGGCCA ACAACAACGT CAACGACCTT CTCTGGACA
31441 TGGACGGCCG CGCCTCGGAG CAGCGACTCG CCAACTTCG CATTGCGCAG CAGCAGGAGA
31501 GAGCCGTCAA GGAGCTGCAG GACGGCATAG CCATCCACCA GTGCAAGAAA GGCATCTTCT
31561 GCCTGGTGAA ACAGGCCAAG ATCTCCTACG AGGTCACCCC GACCGACCAT CGCCTCTCCT
31621 ACGAGCTCCT GCAGCAGCGC CAGAAGTTCA CCTGCCTGGT CGGAGTCAAC CCCATCGTCA
31681 TCACCCAGCA GTCGGGCGAT ACCAAGGGGT GCATCCACTG CTCTGCGGAC TCCCCGACT
31741 GCGTCCACAC TCTGATCAAG ACCCTCTGCG GCCTCCGCGA CCTCTCCCC ATGAACTAAT
31801 CACCCCTTA TCCAGTGAAA TAAATATCAT ATTGATGATG ATTTAAATAA AAAATAATCA
31861 TTTGATTTGA AATAAAGATA CAATCATATT GATGATTTGA GTTTTAAAAA ATAAAGAATC
31921 ACTTACTTGA AATCTGATAC CAGGTCTCTG TCCATGTTTT CTGCCAACAC CACCTCACTC
31981 CCCTCTTCCC AGCTCTGGTA CTGCAGACCC CGGCGGGCTG CAAACTTCCT CCACACGCTG
32041 AAGGGGATGT CAAATTCCTC CTGTCCCTCA ATCTTCATTT TATCTTCTAT CAGATGTCCA
32101 AAAAGCGCGT CCGGGTGGAT GATGACTTCG ACCCGTCTA CCCCTACGAT GCAGACAACG
32161 CACCGACCGT GCCCTTCATC AACCCCCCTC TCGTCTCTTC AGATGGATTG CAAGAGAAGC
32221 CCCTGGGGGT GCTGTCCCTG CGACTGGCTG ACCCGCTCAC CACCAAGAAC GGGGAAATCA
32281 CCCTCAAGCT GGGAGAGGGG GTGGACCTCG ACTCCTCGGG AAAACTCATC TCCAACACGG

```

Fig. 8I

SEQ ID No: 4

69/153

```

32341 CCACCAAGGC CGCCGCCCT CTCAGTTTTT CCAACAACAC CATTTCCTT AACATGGATA
32401 CCCCTCTTTA TACCAAAGAT GGAAAATTAT CCTTACAAGT TTCTCCACCG TTAAACATAT
32461 TAAAATCAAC CATTCTGAAC ACATTAGCTG TAGCTTATGG ATCAGGTTTA GGACTGAGTG
32521 GTGGCACTGC TCTTGCACTA CAGTTGGCCT CTCCACTCAC TTTTGATGAA AAAGGAAATA
32581 TTAAAATTAA CCTAGCCAGT GGTCCATTAA CAGTTGATGC AAGTCGACTT AGTATCAACT
32641 GCAAAAGAGG GGTCACTGTC ACTACCTCAG GAGATGCAAT TGAAAGCAAC ATAAGCTGGC
32701 CTAAAGGTAT AAGATTTGAA GGTAATGGCA TAGCTGCAAA CATTGGCAGA GGATTGGAAT
32761 TTGGAACCAC TAGTACAGAG ACTGATGTCA CAGATGCATA CCCAATTCAA GTTAAATTGG
32821 GTACTGGCCT TACCTTTGAC AGTACAGCGC CCATTGTTGC TTGGAACAAA GAGGATGATA
32881 AACTTACATT ATGGACCACA GCCGACCCCT CGCCAAATTG CAAAATATAC TCTGAAAAAG
32941 ATGCCAACT CACACTTTGC TTGACAAAGT GTGGAAGTCA AATTCTGGGT ACTGTGACTG
33001 TATTGGCAGT GAATAATGGA AGTCTCAACC CAATCACAAA CACAGTAAGC ACTGCCTCG
33061 TCTCCCTCAA GTTTGATGCA AGTGGAGTTT TGCTAAGCAG CTCCACATTA GACAAAAGAT
33121 ATTGGAACCT CAGAAAGGGA GATGTTACAC CTGCTGAGCC CTATACTAAT GCTATAGGTT
33181 TTATGCCTAA CATAAAGGCC TATCCTAAAA ACACATCTGC AGCTTCAAAA AGCCATATTG
33241 TCAGTCAAGT TTATCTCAAT GGGGATGAGG CCAAACCACT GATGCTGATT ATTACTTTTA
33301 ATGAAACTGA GGATGCAACT TGCACCTACA GTATCACTTT TCAATGGAAA TGGGATAGTA
33361 CTAAGTACAC AGGTGAAACA CTTGCTACCA GCTCCTTCAC CTTCTCCTAC ATCGCCCAAG
33421 AATGAACACT GTATCCCACC CTGCATGCCA ACCCTTCCCA CCCCACTCTG TCTATGGAAA
33481 AAACCTCTGA GCACAAAATA AAATAAAGTT CAAGTGTTTT ATTGATTCAA CAGTTTTACA
33541 GGATTCTGAG AGTTATTTTT CCTCCACCCT CCCAGGACAT GGAATACACC ACCCTCTCCC
33601 CCCGCACAGC CTTGAACATC TGAATGCCAT TGGTGATGGA CATGCTTTTG GTCTCCACGT
33661 TCCACACAGT TTCAGAGCGA GCCAGTCTCG GGTCGGTCAG GGAGATGAAA CCTCCGGGC
33721 ACTCCCGCAT CTGCACCTCA CAGCTCAACA GCTGAGGATT GTCCTCGGTG GTCGGGATCA
33781 CGGTTATCTG GAAGAAGCAG AAGAGCGGCG GTGGGAATCA TAGTCCGCGA ACGGGATCGG
33841 CCGGTGGTGT CGCATCAGGC CCCGCAGCAG TCGCTGCCGC CGCCGCTCCG TCAAGCTGCT
33901 GCTCAGGGGG TCCGGGTCCA GGGACTCCCT CAGCATGATG CCCACGGCCC TCAGCATCAG
33961 TCGTCTGGTG CGGCGGGCGC AGCAGCGCAT GCGGATCTCG CTCAGGTCGC TGCAGTACGT
34021 GCAACACAGG ACCACCAGGT TGTTC AACAG TCCATAGTTC AACACGCTCC AGCCGAAACT
34081 CATCGCGGGA AGGATGCTAC CCAGTGCCG GTCTGACCAG ATCCTCAGGT AAATCAAGTG
34141 GCGCTCCCTC CAGAACACGC TGCCACGTA CATGATCTCC TTGGGCATGT GCGGTTTAC
34201 CACCTCCCGG TACCACATCA CCCTCTGGTT GAACATGCAG CCCCAGGATG TCCGTCGGA
34261 CCACAGGGCC AGCACCGCCC CGCCCGCCAT GCAGCGAAGA GACCCCGGGT CCCGGAATG
34321 GCAATGGAGG ACCCACCGCT CGTACCCGTG GATCATCTGG GAGCTGAACA AGTCTATGTT
34381 GGCACAGCAC AGGCATATGC TCATGCATCT CTTCAGCACT CTCAGCTCCT CGGGGTCAA
34441 AACCATATCC CAGGGCACGG GGAACCTCTT CAGGACAGCG AACCCCGCAG AACAGGGCAA
34501 TCCTCGCACA TAACTTACAT TGTGCTAGGA CAGGGTATCG CAATCAGGCA GCACCGGGTG
34561 ATCCTCCACC AGAGAAGCGC GGGTCTCGGT CTCCTCACAG CGTGGTAAGG GGGCCGGCCG
34621 ATACGGGTGA TGGCGGGACG CGGCTGATCG GTTTCGCGAC CGTGTCTATG TGCAGTTCGT
34681 TTCGGACATT TTCGTACTTG CTGTAGCAGA ACCTGGTCCG GCGCTGCAC ACCGATCGCC
34741 GCGGCGGGTC CCGGCGCTTG GAACGCTCGG TGTGAAATT GTAAAACAGC CACTCTCTCA
34801 GACCGTGCAG CAGATCTAGG GCCTCAGGAG TGATGAAGAT CCCATCATGC CTGATAGCTC
34861 TGATCACATC GACCACCGTG GAATGGGCCA GACCCAGCCA GATGATGCAA TTTTGTGGG
34921 TTTCGGTGAC GCGGGGGAG GGAAGAACAG GAAGAACCAT GATTAACTTT TAATCCAAAC
34981 GGTCTCGGAG CACTTCAAAA TGAAGTCCGC GGAGATGGCA CCTCTCGCCC CCGCTGTGTT
35041 GGTGAAAAT AACAGCCAGG TCAAAGGTGA TACGGTTCTC GAGATGTTCC ACGGTGGCTT
35101 CCAGCAAAGC CTCCACGCGC ACATCCAGAA ACAAGACAAT AGCGAAAGCG GGAGGTTCT
35161 CTAATTCCTC AATCATCATG TTACACTCCT GCACCATCCC CAGATAATTT TCATTTTCC
35221 AGCCTTGAAT GATTGCAACT AGTTCCCTGAG GTAAATCCAA GCCAGCCATG ATAAAGAGCT
35281 CGCGCAGAGC GCCCTCCACC GGCATTCTTA AGCACACCCT CATAATTCCA AGATATTCTG
35341 CTCCTGGTTC ACCTGACAGC GATTGACAAG CGGAATATCA AAATCTCTGC CGCGATCCCT
35401 AAGCTCCTCC CTCAGCAATA ACTGTAAAGT CTCTTTCATA TCCTCTCCGA AATTTTAGC
35461 CATAGGACCA CCAGGAATAA GATTAGGGCA AGCCACAGTA CAGATAAACC GAAGTCTCC
35521 CCAGTGAGCA TTGCCAAATG CAAGACTGCT ATAAGCATGC TGGCTAGACC CCGTGATATC
35581 TTCCAGATAA CTGGACAGAA AATCACCCAG GCAATTTTTA AGAAAATCAA CAAAAGAAAA
35641 ATCCTCCAGG TGCACGTTTA GAGCCTCGGG AACAACGATG AAGTAAATGC AAGCGGTGCG
35701 TTCCAGCATG GTTAGTTAGC TGATCTGTAA AAAACAAAAA ATAAACATT AAACCATGCT
35761 AGCCTGGCGA ACAGGTGGGT AAATCGTTCT CTCCAGCACC AGGCAGGCCA CGGGGTCTCC
35821 GCGCGACCC TCGTAAAAAT TGTGCTATG ATTGAAAACC ATCAGAGAGA GACGTTCCCG
35881 GTGGCCGGCG TGAATGATTC GACAAGATGA ATACACCCC GGAACATTGG CGTCCGCGAG

```

Fig. 8J

ITR0048PV

SEQ ID No: 4

70/153

```
35941 TGAAGAAAAAG CGCCCGAGGA AGCAATAAGG CACTACAATG CTCAGTCTCA AGTCCAGCAA
36001 AGCGATGCCA TGCGGATGAA GCACAAAATC CTCAGGTGCG TACAAAATGT AATTACTCCC
36061 CTCCTGCACA GGCAGCGAAG CCCCCGATCC CTCCAGATAC ACATACAAAG CCTCAGCGTC
36121 CATAGCTTAC CGAGCAGCAG CACACAACAG GCGCAAGAGT CAGAGAAAGG CTGAGCTCTA
36181 ACCTGTCCAC CCGCTCTCTG CTCAATATAT AGCCCAGATC TACACTGACG TAAAGGCCAA
36241 AGTCTAAAAA TACCCGCCAA ATAATCACAC ACGCCCAGCA CACGCCCAGA AACCGGTGAC
36301 AACTCAGAAA AAATACGCGC ACTTCCTCAA ACGCCCAAAC TGCCGTCATT TCCGGGTTCC
36361 CACGCTACGT CATCGGAATT CGACTTTCAA ATTCCGTCGA CCGTTAAAAA CGTCACCCGC
36421 CCCGCCCTA ACGGTCGCCC GTCTCTCGGC CAATCACCTT CCTCCCTCCC CAAATTCAAA
36481 CAGCTCATTT GCATATTAAAC GCGCACCAA AGTTTGAGGT ATATTATTGA TGATG
```

Fig. 8K

SEQ ID NO: 5

71/153

```

1 catcatcaat aatataacctc aaacttttgg tgcgcgttaa tatgcaaag aggtatttga
61 atttggggat gcggggcggt gattggctgc gggagcggcg accgttaggg gcggggcggt
121 tgacgtttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt
181 gacgtcaaac gaggtgtggt ttgaacacgg aaataactcaa ttttcccgcg ctctctgaca
241 ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg
301 aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag
361 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat
421 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggt gtcagctgat cgccagggtg
481 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct
541 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg
601 gtaatgtttt cctggctact gggaacgaga ttctgggaact ggtggtggac gccatgatgg
661 gtgacgacct tccggagccc cctaccccc ttgaagcgcc ttcgctgtac gatttgtatg
721 atctggagggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta
781 gcgatgccgc gctgctggct gccgagcagg ctaatacggg ctctggctca gacagcgatt
841 cctctctcca taccgccaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg
901 aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg
961 aggaggcgat tcgagctgca gcgaaccagg gagtgaatac agcgagcgag ggctttagcc
1021 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata
1081 ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt
1141 acagtaagtg tgattaaactt tagctgggga ggcagagggt gactgggtgc tgactgggtt
1201 atttatgtat atgtttttta tgtgtaggtc cgtctctga cgtagatgag acccccacta
1261 cagagtgcac ttcatacccc ccagaaattg gcgaggaacc gcccgagat attattcata
1321 gaccagttgc agtgagagtc accgggcgta gagcagctgt ggagagtttg gatgacttg
1381 tacagggtgg ggatgaacct ttggacttgt gtacccggaa acgccccagg cactaagtgc
1441 cacacatgtg tgtttactta aggtgatgtc agtatttata ggggtgtggag tgcaataaaa
1501 tccgtgttga ctttaagtgc gtggtttatg actcaggggt ggggactgtg ggtatataag
1561 cagggtgcaga cctgtgtggt cagttcagag caggactcat ggagatctgg acagtcttgg
1621 aagactttca ccagactaga cagctgctag agaactcatc ggagggagtc tcttacctgt
1681 ggagattctg cttcgggtggg cctctagcta agctagtcta tagggccaag caggattata
1741 aggatcaatt tgaggatatt ttgagagagt gtccctggat ttttgactct ctcaacttgg
1801 gccatcagtc tcaactttaac cagagtattc tgagagccct tgacttttct actcctggca
1861 gaactaccgc cgcggtagcc ttttttgctt ttatccttga caaatggagt caagaaaccc
1921 atttcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggagggt
1981 gccagcgctt gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga
2041 tcctgagttc ccagtcaccc caggaacacc aacgccgcca gcagccgcag caggagcagc
2101 agcaagagga ggaccgagaa gagaacctga gagccggtct ggaccctccg gtggcggagg
2161 aggaggagta gctgacttgt ttcccgagct gcgcgggttg ctgactaggt ctccagttgg
2221 acgggagagg gggattaagc gggagaggca tgaggagact agccacagaa ctgaactgac
2281 tgtcagttct atgagtcgca ggcgcccaga atcggtgttg tggcatgagg tgcagtcgca
2341 ggggatagat gaggtctcag tgatgcatga gaaatattcc ctagaacaag tcaagacttg
2401 ttggttggag cccgaggatg attgggagg agccatcagg aattatgcca agctggctct
2461 gaggccagac aagaagtaca agattaccaa actgattaat atcagaaatt cctgctacat
2521 ttcaggggat ggggccgagg tggagatcag taccaggag aggtggcct tcagatgctg
2581 catgatgaat atgtaccggg ggtggttgg catggaggga gtcaccttta tgaacggag
2641 gttcaggggt gatgggtata atggggtggt ctttatggcc aacaccaagc tgacagtga
2701 cggatgctcc ttctttggct tcaataacat gtgcatttag gcctggggca gtgtttcagt
2761 gaggggatgc agtttttcag ccaactggat gggggtcgtg ggcagaacca agagcatggt
2821 gtcagtgaag aaatgcctgt tcgagagggt ccacctgggg gtgatgagcg agggcgaagc
2881 caaagtcaaa cactgcgcct ctaccgagac gggctgcttt gtactgatca agggcaatgc
2941 caaagtcaag cataatatga tctgtggggc ctcggtatgag cgcggtacc agatgctgac
3001 ctgcgccggt gggaacagcc atatgctagc caccgtgcat gtggcctcgc acccccga
3061 gacatggccc gagttcgagc acaacgtcat gaccgctgc aatgtgcacc tggggtccc
3121 ccgaggcatg ttcattgccct accagtgcaa catgcaattt gtgaagggtg tgctggagcc
3181 cgatgccatg tccagagtga gcctgacggg ggtgtttgac atgaatgtgg agctgtggaa
3241 aattctgaga tatgatgaat ccaagaccag gtgcccggcc tgcgaatgcg gaggcaagca
3301 cgccaggctt cagcccgtgt gtgtggaggt gacggaggac ctgcgaccgg atcatttgg
3361 gttgtcctgc aacgggacgg agttcggctc cagcggggaa gaatctgact agagtgaata
3421 gtgtttggga ctgggtggga gcctgcatga tgggcagaat gactaaaaatc tgtgttttct

```

Fig. 9A

3481 tgcgcagcag catgagcggg agcgccctct ttgagggagg ggtattcagc ccttatctga
3541 cggggcgtct cccctcctgg gcgggagtcg gtcagaatgt gatgggatcc acgggtggacg
3601 gccggcccgt gcagcccgcg aactcttcaa ccctgacctg cgcgacctg agctcctcgt
3661 ccgtggacgc agctgccgcg gcagctgctg cttccgcccgc cagcgccgtg cgcggaatgg
3721 ccctgggccc cggtacttac agctctcttg tggccaactc gagttccacc aataatcccg
3781 ccagcctgaa cgaggagaag ctgctgctgc tgatggccca gctcgaggcc ctgaccacgc
3841 gcctggtgca ctggaccag caggtggctc agctgcaggc ggagacgcgg gccgcggttg
3901 ccacggtgaa aaccaaataa aaaatgaatc aataaataaa cggagacggt tgttgatttt
3961 aacacagagt cttgaatctt tatttgattt ttcgcgcgcg gtaggccctg gaccaccggt
4021 ctcgatcatt gagcaccggg tggatctttt ccaggaccgg gtagagggtg gcttgatgt
4081 tgaggtagat gggcatgagc ccgtcccggg ggtggaggta gctccattgc agggcctcgt
4141 gctcgggggt ggtgttgtaa atcaccaggt catagcaggg gcgcaggggc tgggtgctga
4201 cgatgtcctt gaggaggaga ctgatggcca cgggcagccc cttggtgtag gtgttgacga
4261 acctgttag agctggagg ctgggaggg gggagatgag atgcattctg gccctggatc
4321 tgagattggc gatgttccc cccagatccc gccgggggtt catgttgtg accaccacca
4381 gcacggtgta tccggtgcac ttggggaatt tgtcatgcaa cttggaaggg aaggcgtgaa
4441 agaatttgga gacgcccctg tgaccgcccc ggttttccat gcactcatcc atgatgatgg
4501 cgatgggccc gtggggcgcg gcttgggcaa agacgtttcg ggggtcggac acatcgtagt
4561 tgtggtcctg ggtgagctcg tcataggcca ttttaatgaa tttggggcgg aggggtgccc
4621 actgggggac gaaggtgccc tcatcccgg gggcgtagtt gccctcgcag atctgcattc
4681 cccaggcctt gagctcggag ggggggatca tgtccacctg cggggcgatg aaaaaaacgg
4741 tttccggggc gggggagatg agctggggcg aaagcagggt ccggagcagc tgggacttgc
4801 cgcagccggt gggggcgtag atgacccgca tgaccggctg cagggtgtag ttgaggaga
4861 gacagctgcc gtccctcgcg aggaggggg ccacctcgtt catcatctcg cgcacatgca
4921 tgttctcgcg cacgagttcc gccaggaggc gctcgcccc aagcgagagg agctcttgca
4981 gcgaggcgaa gtttttcagc ggcttgagcc cgtcggccat gggcattttg gagagggtct
5041 gttgcaagag ttccagacgg tcccagagct cggtagatgt ctctagggca tctcgatcca
5101 gcagacctcc tcgtttcgcg ggttgggcg actgcgggag tagggcacca ggcatgggc
5161 gtccagcgag gccagggtcc ggtccttcca gggcgcgagg gtcgcgtca cgtggtctc
5221 cgtcacggtg aaggggtgcg cgccgggctg ggcgcttgcg aggggtcgct tccgctcat
5281 ccggctggtc gagaaccgct cccggtcgcc gccctcgcg tggccagggt agcaattgag
5341 catgagttcg tagttgagcg cctcggccgc gtggcccttg gcgcggagct tacctttgga
5401 agtgtgtccg cagacgggac agaggaggga cttgagggcg tagagcttgg gggcgaggaa
5461 gacggactcg ggggcgtagg cgtccgcgcc gcagctggcg cagacggtct cgcactccac
5521 gagcgagggt aggtctggcc ggtcggggtc aaaaacgagg ttctctcgt gctttttgat
5581 gcgtttctta cctctggtct ccatgagtc gtgtcccgc tgggtgacaa agaggctgtc
5641 cgtgtccccg tagaccgact ttatgggccc gtccctcagc ggggtgccc cgtcctcgtc
5701 gtagaggaac cccgcccact ccgagacgaa ggcccgggtc caggccagca cgaaggaggc
5761 cacgtgggag gggtagcggg cgttgtccac cagcgggtcc accttctcca gggtagcaa
5821 gcacatgtcc cctcgtcca catccaggaa ggtgattggc ttgtaagtgt aggccacgtg
5881 accgggggtc ccggccgggg gggataaaaa gggggcgggc cctgtctcgt cctcactgtc
5941 ttccggatcg ctgtccagga gcgccagctg ttggggtagg tattccctct cgaaggcggg
6001 catgacctcg gactcagggt tgtcagtttc tagaaacgag gaggatttga tattgacggt
6061 gccgttggag acgcctttca tgagccctc gtccatctgg tcagaaaaga cgtcttttt
6121 gttgtcgagc ttggtggcga aggagccgta gaggcggttg gagagcagct tggcgatgga
6181 gcgcatggtc tggttctttt ccttgtcgcc gcgctccttg gcggcgatgt tgagctgcac
6241 gtactcgcgc gccacgcact tccattcggg gaagacgggt gtgagcttgt cgggcacgat
6301 tctgaccgcg cagccgcggg tgtgcagggt gatgaggtcc acgctgggtg ccacctcgcc
6361 gcgcaggggc tcgttgggtc agcagagggc cccgcccttg cgcgagcaga aggggggcag
6421 cgggtccagc atgagctcgt cgggggggtc ggcgtccacg gtgaagatgc cgggcaggag
6481 ctcggggtcg aagtagctga tgcagggtgc cagatcgtcc agcgccgctt gccagtcgcg
6541 cacggccagc gcgcgctcgt aggggctgag gggcggtgcc cagggcattg ggtgcgtgag
6601 cgcggaggcg tacatgccgc agatgtcgta gacgtagagg ggctcctcga ggacgccgat
6661 gtaggtgggg tagcagcgcc ccccgcggat gctggcgcg acgtagtcgt acagctcgtg
6721 cgagggcgcg agggagccgg tgccgaggtt ggagcgctgc ggcttttcgg cgcggtagac
6781 gatctggcgg aagatggcgt gggagttgga ggagatgggt ggcctctgga agatgttgaa
6841 gtgggcgtgg ggcagtccga ccgagtcctt gatgaagtgg gcgtaggagt cctgcagctt

Fig. 9B

SEQ ID NO: 5

73/153

```

6901 ggcgacgagc tcggcgggtga cgaggacgtc cagggcgag tagtcgaggg tctcttggat
6961 gatgtcgtac ttgagctggc ccttctgctt ccacagctcg cggttgagaa ggaactcttc
7021 gcggtccttc cagtactctt cgagggggaa cccgtcctga tcggcacggt aagagcccac
7081 catgtagaac tggttgacgg ccttgttaggc gcagcagccc ttctccacgg ggagggcgta
7141 agcttgcgcg gccttgcgca gggaggtgtg ggtgagggcg aaggtgtcgc gcaccatgac
7201 cttgaggaac tgggtgcttga agtcagggtc gtgcagccg ccctgtctcc agagctggaa
7261 gtccgtgcgc ttcttgtagg cgggggttgg caaagcgaaa gtaacatcgt tgaagaggat
7321 cttgcccgcg cggggcatga agttgcagat gatcgggaaa ggctggggca cctcgcccgc
7381 gttgttgatg acctgggcgg cgaggacgat ctctcgaaag ccgttgatgt tgtgcccagc
7441 gatgtagagt tccacgaatc gcgggcggcc cttgacgtgg ggcagcttct tgagctcgtc
7501 gtaggtgagc tcggcggggg cgctgaggcc gtgtgtctcg agggcccagt cggcgagggtg
7561 ggggttggcg ccgaggaagg aagtccagag atccacggcc agggcggtct gcaagcggtc
7621 ccggtactga cggaaactgt ggcccagcgc catTTTTTcg ggggtgacgc agtagaaggt
7681 gcgggggtcg ccgtgccagc ggtcccactt gagctggagg gcgaggtcgt gggcgagctc
7741 gacgagcggc ggggtccccg agagtttcat gaccagcatg aaggggacga gctgttgcg
7801 gaaggacccc atccaggtgt aggtttccac gtctgaggtg aggaagagcc tttcgttgcg
7861 aggatgcgag ccgatgggga agaactggat ctctgccac cagttggagg aatggctgtt
7921 gatgtgatgg aagtagaaat gccgacggcg cgccgagcac tcgtgcttgt gttatacaa
7981 gcgtccgcag tgctcgcaac gctgcacggg atgcacgtgc tgcacgagct gtacctgggt
8041 tcctttgacg aggaatttca gtggcgatg gagcgctggc ggctgcatct ggtgtgtac
8101 tacgtcctgg ccatcgccgt ggccatcgtc tgctcgatg gtgtcatgc tgacgaggcc
8161 gcgcgggagg caggtccaga cctcggtcgc gacgggtcgg agagcgagga cgagggcgcg
8221 caggccggag ctgtccaggg tcctgagacg ctgaggagtc aggtcagtgg gcagcgccgg
8281 cgcgcgggtt acttgcaagg gcttttccag ggcgcgcggg aggtccagat ggtacttgat
8341 ctccacggcg ccgttgggtg cgacgtccac ggcttgagg gtcccgtgcc cctggggcgc
8401 caccaccgtg ccccgtttct tcttgggtgc tggcgccggc ggctccatgc ttagaagcgg
8461 cggcgaggac gcgcgcggcg cgccaggggc ggctcggggc ccggaggcag gggcggcagg
8521 ggcacgtcgg cgccgcgcgc gggcagggtt tggtagtcg cccggagaag actggcgtag
8581 gcgacgacgc gacggttgac gtcttgatc tgacgcctct gggtagaagg cacgggaccc
8641 gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtagctgt gacggcgcc
8701 tgccgcagga tctcttgac gtgcgccgag ttgtcctggt aggcgatctc ggtcatgaac
8761 tgctcgatct cctcctcctg aaggtctccg cgaccggcgc gctcgacggg ggccgcgagg
8821 tcgttgagga tgccggccat gagctgcgag aaggcgttca tgccggcctc gttccagacg
8881 cggctgtaga ccacggctcc gtcggggctc gcgcgcgca tgaccacctg ggcgaggttg
8941 agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc
9001 gtggtggcga tgtgtcggg gacgaagaag tacatgatcc agcggcgagg cggcatctcg
9061 ctgacgtcgc ccagggtctc caagcgctcc atggcctcgt agaagtcac ggcgaagtgtg
9121 aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg
9181 gcgatggtgg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc ttccatctcc
9241 tcctcctctt ccatctctc cactaacatc tcttctactt cctcctcagg aggcggcggc
9301 gggggagggg ccctgcgtcg ccggcgcgcg acgggcagac ggtcgatgaa gcgctcgatg
9361 gtctccccgc gccggcgacg catggtctcg gtgacggcgc gccgctctc gcggggccgc
9421 agcgtgaaga cgccgccgcg catctccagg tggccgcggg gggggtctcc gttgggcagg
9481 gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg
9541 agcgtctcga gatccacggg atccgaaaac cgctgaacga aggtctcgag ccagtcgcag
9601 tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcgggcgg
9661 gcgatgctgc tgggtgatga gttgaagtag gcggtcctga gacggcggat ggtggcgagg
9721 agcaccaggt ccttgggccc ggctgtgtg atgcgcagac ggtcgcccat gccccaggcg
9781 tggctcctgac acctggcgag gtcttctgag tagtctgca tgagccgctc cacgggcacc
9841 tctcctcgc cgcgcgggcc gtgcatgcgc gtgagccgca acccgcgctg cggctggacg
9901 agcgccaggt cggcgacgac gcgctcggcg aggatggcct gctggatctg ggtgaggtg
9961 gtctggaagt cgtcgaagtc gacgaagcgg tggtaggctc cgggtgtgat ggtgtaggag
10021 cagttggcca tgacggacca gttgacggtc tggtaggccc ggcgcacgag ctctggttac
10081 ttgaggcgcg agtaggcgcg cgtgtcgaag atgtagtctg tgcaggtgcg cacgaggtac
10141 tggtagtcga cgaggaaagt cgcggcggcg tggcggtaga gcggccatcg ctcggtggcg
10201 ggggcgcggg gcgcgaggtc ctgcagcatg aggcggtggt agccgtagat gtacctggac
10261 atccaggtga tgccggcggc ggtggtggag gcgcgcggga actcgcgagc gcggttccag
10321 atgttgcgca gcggcaggaa gtagttcatg gtggcccgcg tctggcccgt gaggcgcgcg
10381 cagtcgtgga tgctctagac atacgggcaa aaacgaaagc ggtcagcggc tcgactcgt

```

Fig. 9C

10441 ggcttgagg ctaagcgaac gggttgggct gcgcgtgtac cccggttcga gtccctgtct
10501 gaatcaggct ggagccgcag ctaacgtggt actggcactc ccgtctcgac ccaagcctgc
10561 taacgaaacc tccaggatac ggagccgggt cgttttggtc attttcgtca ggccggaat
10621 gaaactagta agcgcgaaa gcggccgtcc gcgatggctc gctgccgtag tctggagaaa
10681 gaatcgccag ggttgcggtg cgtgtgccc cggttcgagc ctcagcgctc ggcgcggcc
10741 ggattccgag gctaacgtgg gcgtggctgc cccgtcgttt ccaagacccc ttagccagcc
10801 gacttctcca gttacggagc gagccccctt ttttcttggt tttttgccag atgcatccc
10861 tactgcggca gatgcgcccc caccctccac cacaaccgcc cctaccgcag cagcagcaac
10921 agccggcgct tctgcccccg ccccgagcag agcagccagc cactaccgcg gcggccgcg
10981 tgagcggagc cggcggttcag tatgacctgg ccttggaaga gggcgagggg ctggcgcgcc
11041 tgggggctgc gtcgcccggg cggcaccgag gcgtgcagat gaaaaggagc gctgcgagg
11101 cctacgtgcc caagcagaac ctgttcagag acaggagcgg cgaggagccc gaggagatgc
11161 gcgcctcccg cttccacgcg gggcgggagc tgcggcgcgg cctggaccga aagcgggtgc
11221 tgaggggacga ggatttcgag gcggacgagc tgacggggat cagccccgcg cgcgcgacg
11281 tggccgcggc caacctgggt acggcgtagc agcagaccgt gaaggaggag agcaacttc
11341 aaaaatcctt caacaaccac gtgcgcagc tgatcgcgcg cgaggagggt accctgggccc
11401 tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa cccacgagc aagccgctga
11461 cggcgagctc gtttctggtg gtgcagcaca gtccgggaca cgagacgttc agggaggcgc
11521 tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga
11581 gcatcggtgt gcaggagcgc gggctgcgcg tgtccgagaa gctggcgccc atcaacttct
11641 cggtgctgag cctgggcaag tactacgcta ggaagatcta caagacccc tacgtgccc
11701 tagacaagga ggtgaagatc gacgggtttt acatgcgcac gaccctgaaa gtgtgcccc
11761 tgagcgacga tctgggggtg taccgcaacg acaggatgca ccgcgcgggt agcgcagcc
11821 gccggcgcca gctgagcgac caggagctga tgcacagcct gcagcgggcc ctgaccgggg
11881 ccgggaccga gggggagagc tactttgaca tgggcgcgga cctgcgctgg cagcctagcc
11941 gccggccctt ggaagctgcc ggcgttccc cctacgtgga ggaggtggac gatgaggagg
12001 aggggggcca gtacctggaa gactgattgc ggcacgtat ttttgctaga tgcagcaaca
12061 gccaccgccc cctcctgac cgcgatgcg ggcggcgctg cagagccagc cgtccggcat
12121 taactcctcg gacgattgga cccaggccat gcaacgcac atggcgctga cgaccgcaa
12181 tcccgaagcc tttagacagc agcctcaggc caaccgactc tcggccatcc tggaggccgt
12241 ggtgccctcg cgctcgaacc ccacgcacga gaaggtgctg gccatcgtga acgcgctggt
12301 ggagaacaag gccatccgag gcgacgaggg cgggctggtg tacaacgcgc tgctggagcg
12361 cgtggccccc gacgaacctg gacgaacctg gaccgcagtg tgaccgacgt
12421 gcgcgagggc gtgtcgcagc gcgagcggtt ccaccgagag tcgaacctgg gctccatggt
12481 ggcgctgaac gccttctga gcacgcagcc cgccaacgtg cccgggggccc aggaggacta
12541 caccaacttc atcagcgcg ctcggctgat ggtggccgag gtgcccaga gcgaggtgta
12601 ccagtcgggg ccggactact tcttcagagc cagtcgccag ggcttgaga ccgtgaacct
12661 gagccaggct ttcaagaact tgcagggact gtggggcggt caggccccgg tcggggaccg
12721 cgcgacgggt tcgagcctgc tgacgcgcaa ctgcgacctg ggctacctgc ttaacctgta
12781 cttcacggac agcggcagcg cgcacgtgga cgagcagacc taccaggaga taccaccagt
12841 ccgcgagggc atcgggcagg aggacccggg caacctggag gccaccctga acttctgct
12901 gagccgcgag ctgggcccag tccgcgcccc gtacgcgctg agcaccgagg aggagcgcat
12961 gaccaaccgg tcgcagaaga tccgcgcccc gttcctgatg caggaggggg ccacgcccag
13021 cctgcgctac gtgcagcaga gcgtggggct gttcctgatg atgtacgccc gcaaccgccc
13081 cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc gcatgaact cggactactt
13141 gttcatcaat aagctgatgg actacttgca tcgggcccgg ggggttctaca cgggagagta
13201 taccaacgcc atcttgaaac cgcacttgca cccgcgcccc ggttgacgac gctgttctc
13261 cgacatgccc gaccccaacg acgggttctt gtgggacgac ggggaccggg ggcgttctc
13321 gccgcgcccc accaccacca ccgtgtggaa gaaagagggc gagggcgcca gcccctccc
13381 ggcgctgtcc ggtgcgcgcg gtgctgcccg ggcgggtgcc cagcagcgag ctgggtcgcc tgacgcggcc
13441 gagcctgccc ttttcgctga acagcgtgcg cagcagcgag cttcgggccc agcgcgagaa
13501 gcgcctgctg ggcgaggagg agtacctgaa cgtactcctg atgagccgct ggaagacgta
13561 gaacttcccc aataacggga tagagacct ggtggacaag accggcgcca cccgtagacg
13621 cgcgcacgag cacagggacg agccccgagc tagcagcagc gaggattccg ccgacgacg
13681 ccagcggcac gacaggcagc ggggtctggt gtgggacgat gaggattccg ccgacgacg
13741 cagcgtgttg gacttgggtg ggagtgggtg tggtaacccg ttcgctcacc tgcgccccg
13801 tatcgggcgc ctgatgtaag aatctgaaaa aataaaagac ggtactcacc aaggccatgg

Fig. 9D

SEQ ID NO: 5

75/153

```

13861 cgaccagcgt gcgtttttct ctgttggttg tagtagtatg atgaggcgcg tgtacccgga
13921 gggctcctcct ccctcgtacg agagcgtgat gcagcaggcg gtggcgggcg cgtatgcagcc
13981 cccgctggag gcgccttacg tgcccccgcg gtacctggcg cctacggagg ggcggaacag
14041 cattcgttac tccgagctgg cacccttgta cgataccacc cggttgtacc tgggtggacaa
14101 caagtcggcg gacatcgctt cgctgaacta ccagaacgac cacagcaact tcctgaccac
14161 cgtggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga
14221 cgagcgctcg cgggtggggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt
14281 gaacgagttc atgtacagca acaagttcaa ggcggggtg atggtctcgc gcaagacccc
14341 caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt ggggtggagt
14401 tgagctgccc gagggcaact tctcggtgac catgaccatc gatctgatga acaacgcat
14461 catcgacaac tacttgccgg tggggcgcca gaacgggggt ctggagagcg acatcggcgt
14521 gaagttcgac acgcgcaact tccggctggg ctgggacccc gtgaccgagc tgggtgatgcc
14581 gggcgtgtac accaacgagg ccttccaccc cgacatcgct ctgctgcccg gctgcggcgt
14641 ggacttcacc gagagccgct tcagcaacct gctgggcatc cgcaagcggc agcccttcca
14701 ggagggtctt cagatcctgt acgagacct ggaggggggc aacatccccg cgtgctgga
14761 cgtggacgcc tacgagaaaa gcaaggagga tagcgccgcc gcgcgaccg cagcgtggc
14821 caccgcctct accgaggtgc ggggcgataa ttttgctagc gccgcgacac tggcgaggc
14881 cgaggcggtt gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga
14941 gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc
15001 ctacaactac ggcgacccc agaagggcgt gcgctcctgg acgctgctca ccacctcgga
15061 cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagaccgggt
15121 cacccttcgc tccacgcgtc aagttagcaa ctaccgggtg gtggcgccg agctcctgcc
15181 cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgcgcctt
15241 caccctcgct acgcacgtct tcaaccgctt ccccgagaa cagatcctcg ttcgcccgcc
15301 cgcgcccacc attaccaccg tcagtgaaaa cgttctctgt ctcacagatc acgggaccct
15361 gccgctgcgc agcagtatcc ggggagttcca gcgctgacc gtcactgacg ccagacgccg
15421 caccctgccc tacgtctaca aggccttggg cgtagtccgc ccgcgcgtcc tctcgagccg
15481 cacccttcta aaaatgtcca ttctcatctc gccagtaat aacaccgggt ggggcctgcy
15541 cgcgcccagc aagatgtacg gaggcgtcgc ccaacgctcc acgcaacacc ccgtgcgcgt
15601 gcgcgggcac ttccgcgctc cctggggcgc cctcaagggc cgcgtgcgt cgcgcaccac
15661 cgtcgacgac gtgatcgacc aggtggtggc cgacgcgcgc aactacacgc ccgcccgcc
15721 gcccgtctcc accgtggacg ccgtcatcga cagcgtgggt gccgacgcgc gccggtacgc
15781 ccgcgccaag agccggcggc ggcgcatcgc ccggcgccac cggagcacc ccgccatgcg
15841 cgcgcgcgca gccttgctgc gcaggccag gcgcacggga cgcagggcca tgcctagggc
15901 ggccagacgc gcggcctccg gcagcagcag cgccggcagg acccgagac gcgcggccac
15961 ggcgggcgcg gcggccatcg ccagcatgct ccgcccgcg cgcggaacg tgtactgggt
16021 gcgcgacgcc gccaccgggt tgccgctgcc cgtgcgcacc cgccccctc gcaactgaag
16081 atgctgactt cgcgatgttg atgtgtccca gcggcgagga ggatgtccaa gcgcaaatc
16141 aaggaagaga tgctccaggt catcgcgctt gagatctacg gcccgcgcc ggtgaaggag
16201 gaaagaaagc cccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg
16261 gacggactgg tggagtgtgt gcgcgagttc gcccccggc ggcgctgca gtggcgcggg
16321 cggaaagtga aaccgggtgt gcgacccggc accacgggtg tcttcacgcc cggcgagcgt
16381 tccggctccg cctccaagcg ctccctacgac gaggtgtacg gggacgagga catcctcgag
16441 caggcgggcg aacgtctggg cgagtttgct tacggcaagc gcagccgccc cgcgccttg
16501 aaagaggagg cgggtgtccat cccgctggac cacggcaacc ccacgccgag cctgaagccg
16561 gtgaccctgc agcaggtgct gcctggtgcy gcgccgcgcc ggggcttcaa gcgcgagggc
16621 ggcgaggatc tgtacccgac catgcagctg atggtgccc aagcgcagaa gctggaggac
16681 gtgctggagc acatgaaggt ggaccccgag gtgcagccc aggtcaaggt gcggcccatc
16741 aagcaggtgg ccccgggcct gggcgtgacg accgtggaca tcaagatccc cagggagccc
16801 atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggaggt gcagacggat
16861 ccctggatgc cggcaccggc ttccaccacc cgccgaagac tccatcatcc ccacgccggg ctaccgccc
16921 ctgctgatgc ccaactacgc gctgcatect tccatcatcc ccacgccggg ctaccgccc
16981 acgcgcttct acccgggcta caccagcagc cgcccgcca agaccaccac ccgcccgcc
17041 cgtcgtcgca cccgccgcag cagcaccgcy acttccgccc ccgccctggt gcggagagt
17101 taccgagcgc ggcgcgagcc tctgacctg ccgcgcgcgc gctaccaccc gagcatcgcc
17161 atttaactac cgctcctac ttgcagatg ggccctcaca tgccgcttcc gcgtcccat
17221 tacgggctac cgaggaagaa agccggccgc tagaaggctg acggggaacg ggctgcgtc
17281 ccatcaccac cggcgggcgc gcgccatcag caagcggtt gggggaggct tcctgcccgc
17341 gctgatgccc atcatcgccg cggcgatcgg ggcgatcccc ggcatagett ccgtggcgg

```

Fig. 9E

SEQ ID NO: 5

76/153

```

17401 gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga
17461 cgctcctggt cctgtgatgt gtgttttttag atggaagaca tcaatttttc gtccctggca
17521 ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcggaac cagccaactg
17581 aacgggggcg ccttcaattg gagcagtcctc tggagcgggc ttaagaattt cgggtccacg
17641 ctcaaaacct atggcaacaa ggcggtggaac agcagcacag ggcaggcgct gagggaaaag
17701 ctgaaagagc agaacttcca gcagaagggt gtcgatggcc tggcctcggg catcaacggg
17761 gtggtgggacc tggccaacca ggccgtgcag aaacagatca acagccgcct ggacgcggtc
17821 ccgcccgcgg ggtccgtgga gatgccccag gtggaggagg agctgcctcc cctggacaag
17881 cgcggcgaca agcgaccgcg tcccgacgcg gaggagacgc tgctgacgca cacggacgag
17941 ccgcccccggt acgaggaggc ggtgaaactg ggtctgcccc ccacgcggcc cgtggcgctt
18001 ctggccaccg ggggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg
18061 cctccgcctg cttcccgcgc ctccacagtg gctaagcccc tgccgcgggt ggccgtcgcg
18121 tcgcgcgccc cccgaggccg ccccaggcg aactggcaga gcactctgaa cagcatcggt
18181 ggtctgggag tcgagagtgt gaagcgcgcg cgctgctatt aaaagacact gtacgcctta
18241 acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agggcgctgt
18301 cgccgagttg caagatggcc accccatcga tgctgcccc gtgggcgtac atgcacatcg
18361 ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagttcgcc cgcgccacag
18421 acacctactt cagtctgggg aacaagttaa ggaacccac ggtggcgccc acgcacgatg
18481 tgaccaccga ccgcagccag cggctgacgc tgcgcttcgt gcccggtggac cgcgaggaca
18541 acacctactc gtacaaagtg cgctacacgc tggcgtgggg cgacaaccgc gtgctggaca
18601 tggccagcac ctactttgac atccgcgggc tgcctggatcg gggccctagc ttcaaacctt
18661 actccggcac cgcttacaac agcctggctc ccaagggagc gccaacact tgccagtgga
18721 catataaagc tgatggtgat actggtacag aaaaaaccta tacatatgga aatgcgcctg
18781 tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc
18841 ccatttatgc agataaaact tatcaaccag agcctcaagt ggggtgatgct gaatggcatg
18901 acatcactgg tactgatgaa aaatatggag gcagagctct caagcctgac accaaaatga
18961 agccctgcta tggttctttt gccaaagccta ccaataaaga aggaggtcag gcaaagtgtg
19021 aaaccgaaac aggcggtacc aaagaatatg acattgacat ggcattcttc gataatcgaa
19081 gtgcagctgc ggctggcctg gccccagaaa ttgttttgta tactgagaat tggtatctgg
19141 aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca
19201 atttgggtca gcagtcctat cccaacagac ccaactacat tggctttaga gacaacttta
19261 tcggggtcat gtactacaac agcactggca acatgggcgt gctggctggt caggcctccc
19321 agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc
19381 ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct
19441 atgaccccg tgtgcgcatt ttgtaaaact acggtgtgga ggatgaactc cctaactatt
19501 gcttccccct ggatgctgtg ggtagaactg atacttacca gggaattaa gccaatggtg
19561 ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaataa ttgggcaagg
19621 gcaatccttt cgccatggag atcaacatcc aggccaaact gtggcggaac ttcctctacg
19681 cgaacgtggc gctgtacctg cccgactcct acaagtacac gccggccaac atcacgctgc
19741 cgaccaacac caacacctac gattacatga acggccgcgt ggtggcgccc tcgctggtgg
19801 acgctacat caacatcggg gcgcgctggt cgctggaccc catggacaac gtcaaccctt
19861 tcaaccacca ccgcaacgcg ggcctgcgt accgctccat gtcctgggg aacgggcgct
19921 acgtgccctt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctctgctcc
19981 tgcccggttc ctacacctac gagtggaaact tccgcaagga cgtcaacatg atcctgcaga
20041 gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc
20101 tctacgccac cttcttcccc atggcgacac acaccgcctc cacgctcgag gccatgctgc
20161 gcaacgacac caacgaccag tccttcaacg actacctctc ggcggccaac atgctctacc
20221 ccattccggc caacgccacc aacgtgcccc tctccatccc ctgcgcgaac tgggcccgtt
20281 tccgcggtg gtcccttcacg cgcttcaaga cccgcgagac gccctcgctc ggtccgggtt
20341 tcgaccccta cttcgtctac tcgggtccca tcccctacct cgacggcacc tctctactca
20401 accacacctt caagaaggte tccatcacct tcgactcctc cgtcagctgg cccggcaacg
20461 accgctcctt gacgccccaa gagttcgaaa tcaagcgcac cgtcgacgga gaggggtaca
20521 acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgctg gccactaca
20581 acatcggtta ccagggtctt tacgtgcccg agggctacaa ggaccgcatg tactccttct
20641 tccgcaactt ccagcccatt agccgcacag tcgtggacga ggtcaactac aaggactacc
20701 aggcgctcac cctggcctac cagcaacaca actcgggctt cgtcggtac ctgcgcccc
20761 ccatgcgcca gggacagccc taccgcgcca actacccta cccgctcctc ggaagagcg

```

Fig. 9F

SEQ ID NO: 5

77/153

```

20821 ccgctcgccag cgtcaccag aaaaagtcc tctgcgaccg ggtcatgtgg cgcacccccct
20881 tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctacg
20941 ccaactccgc ccacgcgcta gacatgaatt tcgaagtcca ccccatggat gagtccaccc
21001 ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg
21061 tcacgcagggc cgtctacctg cgcacgcctt tctcggccgg caacgccacc acctaacccc
21121 cgctcttgct tcttgcaaga tgacggcctg tgcgggctcc ggcgagcagg agctcagggc
21181 catcctccgc gacctgggct ggcgcgcctg cttcctgggc accttcgaca agcgcttccc
21241 gggattcatg gcccgcaca agctggcctg cgccatcgtc aacacggccg gccgcgagac
21301 cgggggcgag cactggctgg ccttcgcctg gaaccgcgc tcccacacct gctacctctt
21361 cgaccccttc ggggttctcg acgagcgctt caagcagatc taccagttcg agtacgagg
21421 cctgctgcgc cgcagcgcgc tggccaccga ggaccgctgc gtcaccttg aaagtccac
21481 ccagaccgtg cagggctcgc gctcggccgc ctgcgggctc ttctgtgca tgttctgca
21541 cgcttcgtg cactggcccg accgccccat ggacaagaac cccaccatga acttgctgac
21601 ggggttgccc aacggcatgc tccagtcgcc ccaggtggaa cccacctgc gccgaacca
21661 ggagcgctc tacgcttcc tcaacgccca ctccgcctac ttctgctccc accgcgcgcg
21721 catcgagaag gccaccgcct tcgaccgcat gaatcaagac atgtaaaccg tgtgtgatg
21781 tgaatgcttt attcataata aacagcacat gtttatgcca ccttttctga ggctctgact
21841 ttatttagaa atcgaagggg ttctgcggcg tctcggcggt ccccgcgggc agggatacgt
21901 tgcggaactg gtacttgggc agccacttga actcggggat cagcagcttc ggcacgggga
21961 ggtcggggaa cgagtcgctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcgg
22021 gcgcggagat cttgaaatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca
22081 cgggggttgc gacatgggac accatcaggg ccgggtgctt cagctcgcgc agcacctcg
22141 cgtcgggtgat gccctccacg tccagatcct cggcggtggc catcccgaag ggggtcatct
22201 tgcaggtctg ccgccccatg ctgggcacgc agccgggctt gtgggtgcaa tcgcagtgca
22261 gggggatcag catcatctgg gcctgctcgg agctcatgcc cgggtacatg gccttcatga
22321 aagcctccag ctggcggaag gcctgctcgc ccttgccgcc ctcggtgaag aagacccgcg
22381 aggacttgct agagaactgg ttggtggcgc agccggcgct gtgcacgcag cagcgcgcgt
22441 cggtgttggc cagctgcacc acgctgcgcc cccagcggtt ctgggtgatc ttggcccggt
22501 cgggggttct cttcagcgcg cgctgcccgt tctcgtcgc cacatccatc tcgatcgtgt
22561 gctccttctg gatcatcag gtcccgtgca ggcacgcag cttgccctcg gcctcgggtg
22621 acccgtgcag ccacagcgcg cagccgggtg actcccagtt cttgtgggag atctgggagt
22681 gcgagtgac gaagccctgc aggaagcggc ccatcatcgt ggtcagggtc ttgttgctgg
22741 tgaaggtcag cgggatgccc cgggtgctct cgttcacata caggtggcag atgcggcggt
22801 acacctcgcc ctgctcgggc atcagctgga aggcggactt caggtcgctc tccacgcggt
22861 accggtccat cagcagcgtc atgacttcca tgcccttctc ccaggccgag acgatcggca
22921 ggctcagggg gttcttcacc gccgtgtgca tcttagtcgc cgccgtgag gtcagggggg
22981 cgttctcgtc caggggtctc aacactcgct tgccgtcctt ctcggtgatg cgcacggggg
23041 gaaagctgaa gccacgggcc gccagctcct cctcggcctg ccttctgctc tcgctgctc
23101 ggctgatgtc ttgcaaaggc acatgcttgg tcttgcgggg tttctttttg ggccgagag
23161 gcggcgccgg agacgtgctg ggcgagcgcg agttctcgt caccacgact atttcttctt
23221 cttggccgct gtccgagacc acgcggcggt aggcagcct cttctggggc agagggcgag
23281 gcgacgggct ctcgcggttc ggcggcgccg tggcagagcc ccttcgcgt tcgggggtgc
23341 gctcctggcg gcgctgctc gactgacttc ctccgcggcc ggccattgtg ttctcctagg
23401 gagcaacaag catggagact cagccatcgt cgccaacatc gccatctgcc cccgcgcgcg
23461 ccgacgagaa ccagcagcag aatgaaagct taaccgcccc gcgccccagc cccacctccg
23521 acgcccgcgc ggcccagac atgcaagaga tggaggaatc catcgagatt gacctgggct
23581 acgtgacgcc cgcggagcac gaggaggagc tggcagcgcg ctttccagcc ccggaagaga
23641 accaccaaga gcagccagag caggaagcag agagcgagca gcagcaggct gggctcgagc
23701 atggcgacta cctgagcggg gcagaggacg tgctcatcaa gcatctggcc cgccaatgca
23761 tcacttgcaa ggacgcgctg ctcgaccgcg ccgagtgccc cctcagcgtg gcgagctca
23821 gccgcgccta cgagcgcaac ctcttctcgc cgcgcgtgcc ccccaagcgc cagcccaacg
23881 gcacctgcga gcccaacccg cgctcaact tctaccgggt cttcgcggtg cccgagccc
23941 tggccaccta ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc
24001 gcacccgcgc cgacgcctg ctcaacctgg gtcccggcgc ccgctacct gatatgcct
24061 ccttgggaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggcccgcga
24121 acgctctgca aggaagcgga gaggagcatg agcaccacag cgccctgggt gagttggaag
24181 gcgacaacgc gcgctggcg gtgctcaagc gcacggtcga gctgacccac ttgcctacc
24241 cggcgctcaa cctgcccccc aaggtcaatg gcgcgctcat ggaccaggtg ctcacaaagc
24301 gcgctcgcgc cctctcggat gaggacatgc aggaccccga gagctcggac gagggaagc

```

Fig. 9G

24361 ccgtgggtcag cgacgagcag ctggcgcgct ggctggggagc gagtagcacc cccagagct
24421 tggaagagcg ggcgaagctc atgatggcgg ttgtcctggt gaccgtggag ctggagtgtc
24481 tgcgcgcgtt cttcgccgac gcagagaccc tgcgcaaggt cgaggagaac ctgcactacc
24541 tcttcaggca cgggtttgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc
24601 ttgtctccta catgggcatc ctgcacgaga accgcctggg gcagaacgtg ctgcacacca
24661 ccctgcgcgg ggaggcccg cgcgactaca tccgcgactg cgtctacctg tacctctgcc
24721 acacctggca gacgggcatg ggctgtgtgc agcagtgcct ggaggagcag aacctgaaag
24781 agctctgcaa gctcctgcag aagaacctga agggcctgtg gaccgggttc gacgagcgca
24841 ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcggctg acgtgcgca
24901 acggactgcc cgactttatg agtcaaagca tgttgcaaaa ctttcgctct ttcacctctg
24961 aacgctccgg gatcctgccc gccacctgct ccgcgctgcc ctcgacttc gtccgctga
25021 ccttcgcgga gtgcccccg ccgctctgga gccactgcta cctgctgccc ctggccaaact
25081 acctggccta ccactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc
25141 actgccgctg caacctctgc acgcccgcacc gctccctggc ctgcaacccc cagctgctga
25201 gcgagaccca gatcatcggc accttcgagt tgcaaggccc cggcgagggc aagggggggtc
25261 tgaaactcac cccggggctg tggacctcgg cctacttgcg caagtctgtg cccgaggact
25321 accatccctt cgagatcagg ttctacgagg accaatccca gccgccaag gccgaactgt
25381 cggcctgcgt catcaccag ggggccatcc tggcccaatt gcaagccatc cagaaatccc
25441 gccagaatt tctgctgaaa aagggccacg gggctctacct ggacccccag accggagagg
25501 agctcaaccc cagcttcccc caggatgccc cgaggaagca gcaagaagct gaaagtggag
25561 ctgccgcgcg cggaggattt ggaggaagac tgggagagca gtcaggcaga ggaggaggag
25621 atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac
25681 gaggtggagg aggaggcaga ggaagaagca gccgcgcgca gaccgtcgtc ctcgccggag
25741 aaagcaagca gcacggatag catctccgct ccgggtcggg gtcgcggcga ccgggcccac
25801 agtaggtggg acgagaccgg gcgcttcccc aacccacca cccagaccgg taagaaggag
25861 cggcagggat acaagtccct gcgggggacg caccgcgcgc tacctgctct tccaccgagg ggtgaacttc
25921 tgcgggggca acatctcctt caccgcgcgc ctccacagcc cctactactg ttccaagaa
25981 ccccgcaaca tcttgcaata ctaccgtcac ctccacagcc cctactactg ttccaagaa
26041 gaggcagaaa cccagcagca gcagaaaacc agcggcagca gcagctagaa aatccacagc
26101 ggcggcaggt ggaactgagga tcgcagcgaa cgagccggcg cagaccgggg agctgaggaa
26161 ccggatcttt cccaccctct atgccatctt ccagcagagt cgggggcagg agcaggaaact
26221 gaaagtcaag aaccgttctc tgcgctcgct caccgcgagt tgtctgtatc acaagagcga
26281 agaccaactt cagcgactc tcgaggacgc cgaggctctc ttcaacaagt actgcgcgt
26341 cactcttaaa gagtagcccg cgcccgccca cacacggaaa aaggcgggaa ttacgtcacc
26401 acctgcgccc ttgcggcgac catcatcatg agcaaaagaga ttcccacgce ttacatgtgg
26461 agtaccagc cccagatggg cctggccgccc ggcgcgccc aggactactc caccgcgatg
26521 aactggctca gcgcggggcc cgcgatgatc tcacgggtga atgacatccg cgcccgccga
26581 aaccagatac tcctagaaca gtcagcgatc accgccacgc cccgccatca ccttaatccg
26641 cgtaattggc ccgcccgcct ggtgtaccag gaaattcccc agcccacgac cgtactactt
26701 ccgcgagacg cccaggccga agtcacgctg actaactcag gtgtccagct ggccggcgcc
26761 gccgccctgt gtcgtcaccg ccccgctcag ggtataaagc ggctgggtgat ccgaggcaga
26821 ggcacacagc tcaacgacga ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc
26881 ttccaactcg ccggatcggg gagatcttcc ttacgcctc gtcaggccgt cctgactttg
26941 gagagtctgt cctcgagcc cgcctcgggt ggcacggca ctctccagtt cgtggaggag
27001 ttcaactcct cggctactt caaccccttc tccggctccc ccggccacta cccggacgag
27061 ttcatcccga acttcgacgc catcagcgag tcggcttcga cacctggacc actgtcgcct ctcctacgag
27121 ggtggcgag ctgacctagc tccgcttcga cactgggag ctggctcggag tcaacccat cgtcatcacc
27181 ctctgcagc agcgccagaa gttcacctgc cactgctcgt gcgactcccc cgactgcgtc
27241 cagcagtcgg gcgataccaa ggggtgcatc cgtgcacctc tccccatgaa ctaatcacc
27301 cacactctga tcaagaccct ctgcggcctc cgcgacctcc tccccatgaa ctaatcacc
27361 ccttatccag tgaaataaag atcatattga tgatttgagt ttaataaaaa taaagaatca
27421 cttacttgaa atctgatacc aggtctctgt ccatgttttc tgccaacacc acttcactcc
27481 cctcttccca gctctggtac tgcaggcccc ggccgggctgc aaacttcctc cacaccctga
27541 aggggatgtc aaattcctcc tgtccctcaa tcttcatttt atcttctatc agatgtccaa
27601 aaagcgcgtc cgggtggatg atgacttcca cccgctctac ccctacgatg cagacaacgc
27661 accgaccgtg cccttcatca accccccctt cgtctcttca gatggattcc aagagaagcc
27721 cctgggggtg ctgtccctgc gtctggccga tcccgctacc accaagaacy gggaaatcac

Fig. 9H

SEQ ID NO: 5

79/153

27781 cctcaagctg ggagatgggg tggacctcga ctccctcgga aaactcatct ccaacacggc
27841 caccaaggcc gccgcccctc tcagtttttc caacaacacc atttccctta acatggatac
27901 cccttttttac aacaacaatg gaaagttagg catgaaagtc actgctccac tgaagatact
27961 agacacagac ttgctaaaaa cacttggtgt agcttatgga caaggttttag gaacaaacac
28021 cactggtgcc cttggtgccc aactagcatc cccacttgct tttgatagca atagcaaaat
28081 tgcccttaat ttaggcaatg gaccattgaa agtggatgca aatagactga acatcaattg
28141 caatagagga ctctaatgta ctaccacaaa agatgcaact gaagccaata taagttgggc
28201 taatgctatg acatttatag gaaatgccat ggggtgtcaat attgatacac aaaaaggctt
28261 gcaatttgcc accactagta ccgtcgcaga tgtaaaaaac gcttaccoca taaaaatcaa
28321 acttgagact ggtctcacat ttgacagcac aggtgcaatt gttgcatgga acaaagatga
28381 tgacaagctt acactatgga ccacagccga cccctctcca aattgtcaca tatattctga
28441 aaaggatgct aagcttacac tttgcttgac aaagtgtggc agtcagattc tgggcactgt
28501 ttccctcata gctgttgata ctggcagttt aaatcccata acaggaacag taaccactgc
28561 tcttgcttca cttaaattcg atgcaattgg agttttgcaa agcagctcaa cactagactc
28621 agactattgg aatttcagac agggagatgt tacacctgct gaagcctata ctaatgctat
28681 aggtttcatg cccaatctaa aagcataccc taaaaacaca agtggagctg caaaaagtca
28741 cattgttggg aaagtgtacc tacatgggga tacaggcaaa ccaactggacc tcattattac
28801 tttcaatgaa acaagtgatg aatcttgcac ttactgtatt aactttcaat ggcagtgggg
28861 ggctgatcaa tataaaaaatg aaacacttgc cgtcagttca ttcacctttt ccttatattgc
28921 taaagaataa accccactct gtaccccatc tctgtctatg gaaaaaactc tgaaacacaa
28981 aataaaataa agttcaagtg ttttattgat tcaacagttt tacaggattc gagcagttat
29041 ttttctccca cctcccagg acatggaata caccaccctc tcccccgca cagccttgaa
29101 catctgaatg ccattggtga tggacatgct tttggtctcc acgttccaca cagtttcaga
29161 gcgagccagt ctccgggtcg tccaggagat gaaaccctcc gggcactccc gcatctgcac
29221 ctacagctc aacagctgag gattgtcctc ggtggtcggg atcacggtta tctggaagaa
29281 gcagaagagc ggcgggtggga atcatagtcc gcgaacggga tcggccgggtg gtgtcgcac
29341 aggccccgca gcagtcgctg tcgcccggc tccgtcaagc tgctgtctag ggggtccggg
29401 tccagggact cctcagcat gatcccag cctcgtcagg tcgctgcagt acgtgcaaca caggaccacc
29461 gcgcagcagc gcatgaggat ctgctcagg gttcaacacg ctccagccga aactcatcgc gggaaggatg
29521 aggttggtca acagtccata gttcaacacg ctccagccga aactcatcgc gggaaggatg
29581 ctaccacagt ggccgtcgta ccagatcctc aggtaaatca agtggcgccc cctccagaac
29641 acgctgcccc tgtagatgat ctccctgggc atgtggcggt tcaccacctc ccggtaccac
29701 atcacccctc ggttgaacat gcagccccgg atgatcctgc ggaaccacag ggccagcacc
29761 gccccgcccg ccattgcagc aagagacccc ggggtcccgac aatggcaatg gaggaccac
29821 cgctcgtaac cgtggatcat ctgggagctg aacaagtcta tgttggcaca gcacaggcat
29881 atgctcatgc atctcttcag cactctcagc tctcggggg tcaaaaccat atcccagggc
29941 acggggaact cttgcaggac agcgaacccc gcagaacagg gcaatcctcg cacataactt
30001 acattgtgca tggacagggt atcgcaatca ggcagcaccg ggtgatcctc caccagagaa
30061 gcgcggtct cgggtctctc acagcgtggt aagggggccg gccgatacgg gtgatggcg
30121 gacgcggctg atcgtgttcg cgaccgtgtt atgatgcagt tgctttcgga cattttcgta
30181 cttgctgtag cagaacctgg tccgggcgct gcacaccgat cgccggcggc ggtcccggcg
30241 cttggaacgc tcggtgttga agttgtaaaa cagccactct ctgagaccgt gcagcagatc
30301 tagggcctca ggagtgatga agatcccatc atgctgatg gctctaatca catcgaccac
30361 cgtggaatgg gccagaccca gccagatgat gcaattttgt tgggtttcgg tgacggcggg
30421 ggaggggaaga acaggaagaa ccatgattaa cttttaatcc aaacggtctc ggagcacttc
30481 aaaatgaaga tcgaggagat ggcacctctc gcccccgtg tgttgggtga aaataacagc
30541 caggtcaaag gtgatacgggt tctcgagatg ttccacggtg gcttccagca aagcctccac
30601 gcgcacatcc agaacaaga caatagcgaa agcgggaggg ttctctaatt cctcaatcat
30661 catgttacac tccgtcacca tcccagata attttcattt ttccagcctt gaatgattcg
30721 aactagttcc tgaggtaaat ccaagcagc catgataaag agctcgcgca gaggccctc
30781 caccggcatt cttaagcaca cctcataat tccaagatat tctgctctg gttccactgc
30841 agcagattga caagcggaat atcaaaatct ctgccgcgat ccctaagctc ctccctcagc
30901 aataactgta agtactcttt catatcctct ccgaaatctt tagccatagg accaccagga
30961 ataagattag ggcaagccac agtacagata aaccgaagtc ctcccagtg agcattgcca
31021 aatgcaagac tgctataagc atgctggcta gaccgggtga tatcttccag ataactggac
31081 agaaaatcgc ccaggcaatt ttttaagaaa tcaacaaaag aaaaatcctc caggtgcacg

Fig. 9I

ITR0048PV

SEQ ID NO: 5

80/153

```
31141 tttagagcct cgggaacaac gatggagtaa atgcaagcgg tgcgttccag catgggtagt
31201 tagctgatct gtagaaaaaa acaaaaatga acattaaacc atgctagcct ggcgaacagg
31261 tgggttaaate gttctctcca gcaccaggca ggccacgggg tctccggcac gaccctcgta
31321 aaaattgtcg ctatgattga aaaccatcac agagagacgt tcccgggtggc cggcgtgaat
31381 gattcgacaa gatgaataca cccccggaac attggcgctcc gcgagtgaat aaaagcgccc
31441 aaggaagcaa taaggcacta caatgctcag tctcaagtcc agcaaagcga tgccatgcgg
31501 atgaagcaca aaattctcag gtgcgtacaa aatgtaatta ctcccctcct gcacaggcag
31561 caaagcccc gatccctcca ggtacacata caaagcctca gcgtccatag cttaccgagc
31621 agcagcacac aacaggcgca agagtccagag aaaggctgag ctctaacctg tccaccgct
31681 ctctgctcaa tatatagccc agatctacac tgacgtaaag gccaaagtct aaaaaataccc
31741 gccaaataat cacacacgcc cagcacacgc ccagaaaccg gtgacacact caaaaaata
31801 cgcgcacttc ctcaaagccc caaactgccg tcatttccgg gttcccacgc tacgtcatca
31861 aaattcgact ttcaaattcc gtcgaccggt aaaaacgtcg cccgccccgc ccctaacggt
31921 cgcgcgtccc gcagccaatc accgccccgc atccccaaat tcaaatacct catttgcata
31981 ttaacgcgca ccaaagttt gaggtatatt attgatgatg
```

Fig. 9J

ITR0048PV

SEQ ID NO: 6

81/153

1 ATGAAGCGCA CAAAACGTC TGACGAGAGC TTCAACCCCG TGTACCCCTA
TGACACGGAA

61 AGCGGCCCTC CCTCCGTCCC TTTCTCACC CCTCCCTTCG TGTCTCCCGA
TGGATTCCAA

121 GAAAGTCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC
TCCCCACGGC

181 ATGCTCGCCC TGAAAATGGG AAGTGGCCTC TCCCTGGACG ACGCTGGCAA
CCTCACCTCT

241 CAAGATATCA CCACCGCTAG CCCTCCCCTC AAAAAACCA AGACCAACCT
CAGCCTAGAA

301 ACCTCATCCC CCCTAACTGT GAGCACCTCA GCGGCCCTCA CCGTAGCAGC
CGCCGCTCCC

361 CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGAC
AGTACAGGAT

421 GCAAACTCA CCCTGGCCAC CAAAGGCCCC CTGACCGTGT CTGAAGGCAA
ACTGGCCTTG

481 CAAACATCGG CCCCCTGAC GGCCGCTGAC AGCAGCACCC TCACAGTCAG
TGCCACACCA

541 CCCCTTAGCA CAAGCAATGG CAGCTTGGGT ATTGACATGC AAGCCCCCAT
TTACACCACC

601 AATGGAAAAC TAGGACTTAA CTTTGGCGCT CCCCTGCATG TGGTAGACAG
CCTAAATGCA

661 CTGACTGTAG TTAGTGGCCA AGGTCTTACG ATAAACGGAA CAGCCCTACA
AACTAGAGTC

721 TCAGGTGCCC TCAACTATGA CACATCAGGA AACCTAGAAT TGAGAGCTGC
AGGGGGTATG

781 CGAGTTGATG CAAATGGTCA ACTTATCCTT GATGTAGCTT ACCCATTTGA
TGCACAAAAC

841 AATCTCAGCC TTAGGCTTGG ACAGGGACCC CTGTTTGTTA ACTCTGCCCA
CAACTTGGAT

901 GTTAACTACA ACAGAGGCCT CTACCTGTTC ACATCTGGAA ATACCAAAAA
GCTAGAAGTT

961 AATATCAAAA CAGCCAAGGG TCTCATTTAT GATGACACTG CTATAGCAAT
CAATGCGGGT

1021 GATGGGCTAC AGTTTGACTC AGGCTCAGAT ACAAATCCAT TAAAACTAA
ACTTGGATTA

1081 GGAAGTGGATT ATGACTCCAG CAGAGCCATA ATTGCTAAAC TGGGAAGTGG
CCTAAGCTTT

Fig. 10A

ITR0048PV

SEQ ID NO: 6

82/153

1141 GACAACACAG GTGCCATCAC AGTAGGCAAC AAAAATGATG ACAAGCTCAC
CTTGTGGACC

1201 ACACCAGACC CATCTCCTAA CTGTAGAATC TATTCAGAGA AAGATGCTAA
ATTCACACTT

1261 GTTTTGACTA AATGCGGCAG TCAGGTGTTG GCCAGCGTTT CTGTTTTATC
TGTAAGAGGT

1321 AGCCTTGCGC CCATCAGTGG CACAGTAACT AGTGCTCAGA TTGTCTCAG
ATTTGATGAA

1381 AATGGAGTTC TACTAAGCAA TTCTTCCCTT GACCCTCAAT ACTGGAAC TA
CAGAAAAGGT

1441 GACCTTACAG AGGGCACTGC ATATACCAAC GCAGTGGGAT TTATGCCCAA
CCTCACAGCA

1501 TACCCAAAAA CACAGAGCCA AACTGCTAAA AGCAACATTG TAAGTCAGGT
TTACTTGAAT

1561 GGGGACAAAT CCAAACCCAT GACCCTCACC ATTACCCTCA ATGGAAC TA
TGAAACAGGA

1621 GATGCCACAG TAAGCACTTA CTCCATGTCA TTCTCATGGA ACTGGAATGG
AAGTAATTAC

1681 ATTAATGAAA CGTTCCAAAC CAACTCCTTC ACCTTCTCCT ACATCGCCCA
AGAATAA

Fig. 10B

ITR0048PV

SEQ ID NO: 7

83/153

1 ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC
CTACGATGCA

61 GACAACGCAC CGACCGTGCC CTTTCATCAAC CCCCCCTTCG TCTCTTCAGA
TGGATTCCAA

121 GAGAAGCCCC TGGGGGTGTT GTCCCTGCGA CTGGCCGACC CCGTCACCAC
CAAGAACGGG

181 GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGATT CCTCGGGAAA
ACTCATCTCC

241 AACACGGCCA CCAAGGCCGC CGCCCCCTCTC AGTTTTTCCA ACAACACCAT
TCCCCTTAAC

301 ATGGATCACC CCTTTTACAC TAAAGATGGA AAATTATCCT TACAAGTTTC
TCCACCATTA

361 AATATACTGA GAACAAGCAT TCTAAACACA CTAGCTTTAG GTTTTGGATC
AGGTTTAGGA

421 CTCCGTGGCT CTGCCTTGGC AGTACAGTTA GTCTCTCCAC TTACATTGGA
TACTGATGGA

481 AACATAAAGC TTACCTTAGA CAGAGGTTTG CATGTTACAA CAGGAGATGC
AATTGAAAGC

541 AACATAAGCT GGGCTAAAGG TTTAAAATTT GAAGATGGAG CCATAGCAAC
CAACATTGGA

601 AATGGGTTAG AGTTTGGAAG CAGTAGTACA GAAACAGGTG TTGATGATGC
TTACCCAATC

661 CAAGTTAAAC TTGGATCTGG CCTTAGCTTT GACAGTACAG GAGCCATAAT
GGCTGGTAAC

721 AAAGAAGACG ATAAACTCAC TTTGTGGACA ACACCTGATC CATCACCAAA
CTGTCAAATA

781 CTCGCAGAAA ATGATGCAAA ACTAACACTT TGCTTGACTA AATGTGGTAG
TCAAATACTG

841 GCCACTGTGT CAGTCTTAGT TGTAGGAAGT GGAAACCTAA ACCCCATTAC
TGGCACCGTA

901 AGCAGTGCTC AGGTGTTTCT ACGTTTTGAT GCAAACGGTG TTCTTTTAAC
AGAACATTCT

961 AACTAAAAA AATACTGGGG GTATAGGCAG GGAGATAGCA TAGATGGCAC
TCCATATACC

1021 AATGCTGTAG GATTCATGCC CAATTTAAA GCTTATCCA AGTCACAAAG
TTCTACTACT

1081 AAAAATAATA TAGTAGGGCA AGTATACATG AATGGAGATG TTTCAAAACC
TATGCTTCTC

Fig. 11A

ITR0048PV

SEQ ID NO: 7

84/153

1141 ACTATAACCC TCAATGGTAC TGATGACAGC AACAGTACAT ATTCAATGTC
ATTTTCATAC

1201 ACCTGGACTA ATGGAAGCTA TGTTGGAGCA ACATTTGGGG CTAACCTCTTA
TACCTTCTCA

1261 TACATCGCCC AAGAATGA

Fig. 11B

ITR0048PV

SEQ ID NO: 8

85/153

1 ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC
CTACGATGCA

61 GACAACGCAC CGACCGTGCC CTTTCATCAAC CCCCCCTTCG TCTCTTCAGA
TGGATTCCAA

121 GAGAAGCCCC TGGGGGTGCT GTCCCTGCGT CTGGCCGATC CCGTCACCAC
CAAGAACGGG

181 GAAATCACCC TCAAGCTGGG AGATGGGGTG GACCTCGACG ACTCGGGAAA
ACTCATCTCC

241 AACACGGCCA CCAAGGCCGC CGCCCTCTC AGTTTTTCCA ACAACACCAT
TTCCCTTAAC

301 ATGGATACCC CTCTTTACAA CAACAATGGA AAGCTAGGTA TGAAGGTAAC
CGCACCATTA

361 AAGATATTAG ACACAGATCT ACTAAAAACA CTTGTTGTTG CTTATGGGCA
GGGATTAGGA

421 ACAAACACCA ATGGTGCTCT TGTGCCCCAA CTAGCATACC CACTTGTTTT
TAATACCGCT

481 AGCAAAATTG CCCTTAATTT AGGCAATGGA CCATTAAAAG TGGATGCAAA
TAGACTGAAC

541 ATTAATTGCA AAAGAGGTAT CTATGTCACT ACCACAAAAG ATGCACTGGA
GATTAATATC

601 AGTTGGGCAA ATGCTATGAC ATTTATAGGA AATGCCATTG GTGTCAATAT
TGACACAAAA

661 AAAGGCCTAC AGTTCGGCAC TTCAAGCACT GAAACAGATG TTAAAAATGC
TTTTTCACTC

721 CAAGTAAAAC TTGGAGCTGG TCTTACATTT GACAGCACAG GTGCCATTGT
TGCTTGGAAC

781 AAAGAAGATG ACAAACCTAC ACTGTGGACC ACAGCCGATC CATCTCCAAA
CTGTACATA

841 TATTCTGCAA AGGATGCTAA GCTTACACTC TGCTTGACAA AGTGTGCTAG
TCAAATCCTA

901 GGCAGTGTCT CCCTATTAGC AGTCAGTGGC AGCTTGGCTC CTATCACAGG
GGCTGTTAGA

961 ACTGCACTTG TATCACTCAA ATTCAATGCT AATGGAGCCC TTTTGGACAA
ATCAACTCTG

1021 AACAAAGAAT ACTGGAAC TA CAGACAAGGA GATCTAATTC CAGGTACACC
ATATACACAT

1081 GCTGTGGGTT TCATGCCTAA CAAAAAGCC TACCCTAAAA ACACAACTGC
AGCTTCCAAG

Fig. 12A

ITR0048PV

SEQ ID NO: 8

86/153

1141 AGCCACATTG TGGGTGATGT GTATTTAGAT GGAGATGCAG ATAAACCTTT
ATCTCTTATC

1201 ATCACTTTCA ATGAAACTGA TGATGAAACC TGTGATTACT GCATCAACTT
TCAATGGAAA

1261 TGGGGAGCTG ATCAATATAA GGATAAGACA CTCGCAACCA GTTCATTAC
CTTCTCATA

1321 ATCGCCCAAG AATAA

Fig. 12B

ITR0048PV

SEQ ID NO: 9

87/153

1 ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC
CTACGATGCA

61 GACAACGCAC CGACCGTGCC CTTTCATCAAC CCCCCCTTCG TCTCTTCAGA
TGGATTCCAA

121 GAGAAGCCCC TGGGGGTGCT GTCCCTGCGA CTGGCCGACC CCGTCACCAC
CAAGAACGGG

181 GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA
ACTCATCTCC

241 AACACGGCCA CCAAGGCCGC CGCCCCTCTC AGTTTTTCCA ACAACACCAT
TTCCCTTAAC

301 ATGGATACCC CTTTTTACAA CAATAATGGA AAGTTAGGCA TGAAAGTCAC
TGCTCCACTG

361 AAGATACTCG ACACAGACTT GCTAAAAACA CTTGTTGTAG CTTATGGACA
AGGTTTAGGA

421 ACAAAACCCA CTGGTGCCCT TGTGCCCCAA CTAGCAGCCC CACTTGCTTT
TGATAGCAAT

481 AGCAAAATTG CCCTTAATTT AGGCAATGGA CCATTGAAAG TGGATGCAAA
TAGACTGAAC

541 ATCAATTGCA ATAGAGGACT CTATGTTACT ACCACAAAAG ATGCACTGGA
AACCAACATA

601 AGTTGGGCTA ATGCTATGAC ATTTATAGGA AATGCCATGG GTGTCAATAT
TGATACACAA

661 AAAGGCTTGC AATTTGGCAC CACTAGTACC GTCGCAGATG TTAAAAACGC
TTACCCCAT

721 CAAGTCAAAC TGGGAGCTGG TCTCACATTT GACAGCACAG GTGCAATTGT
CGCTTGGAAC

781 AAAGAAGATG ACAAACTTAC ACTGTGGACC ACAGCCGATC CATCTCCAAA
CTGTCACATA

841 TATTCTGACA AGGATGCTAA GCTTACACTC TGCTTGACAA AGTGTGGCAG
TCAGATACTG

901 GGCAGTGTTC CTCTCATAGC TGTGATACT GGTAGCTTAA ATCCAATAAC
AGGACAAGTA

961 ACCACTGCTC TTGTTTCACT TAAATTCGAT GCCAATGGAG TTTTGCAAAC
CAGTTCAACA

1021 TTGGACAAAG AATATTGGAA TTTTAGAAAA GGAGATGTGA CACCTGCTGA
GCCATATACT

1081 AATGCTATAG GTTTCATGCC CAATCTAAAG GCATACCCTA AAAACACAAG
TGGAGCTGCA

Fig. 13A

ITR0048PV

SEQ ID NO: 9

88/153

1141 AAAAGTCACA TTGTTGGGAA AGTGTACCTA CATGGGGATA CAGACAAACC
ACTGGACCTG

1201 ATTATTACTT TCAATGAAAC AAGTGATGAA TCTTGCACTT ACTGTATTAA
CTTTC AATGG

1261 AAATGGGATA GTACTAAGTA CACAGGTGAA ACACTTGCTA CAAGCTCCTT
CACCTTCTCC

1321 TACATTGCCC AAGAATGA

Fig. 13B

ITR0048PV

SEQ ID NO: 10

89/153

1 ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
61 GACAACGCAC CGACCGTGCC CTTCAATCAAC CCCCCCTTCG TCTCTTCAGA
TGGATTCCAA
121 GAGAAGCCCC TGGGGGTGTT GTCCCTGCCA CTGGCCGACC CCGTCACCAC
CAAGAACGGG
181 GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA
ACTCATCTCC
241 AACACGGCCA CCAAGGCCGC TGCCCTCTC AGTTTTTCCA ACAACACCAT
TTCCCTTAAC
301 ATGGATCACC CCTTTTACAC TAAAGATGGA AAATTAGCCT TACAAGTTTC
TCCACCATTA
361 AATATACTGA GAACAAGCAT TCTAAACACA CTAGCTTTAG GTTTTGGATC
AGGTTTAGGA
421 CTCCGTGGCT CTGCCTTGGC AGTACAGTTA GTCTCTCCAC TTACATTTGA
TACTGATGGA
481 AACATAAAGC TTACCTTAGA CAGAGGTTTG CATGTTACAA CAGGAGATGC
AATTGAAAGC
541 AACATAAGCT GGGCTAAAGG TTTAAAATTT GAAGATGGAG CCATAGCAAC
CAACATTGGA
601 AATGGGTTAG AGTTTGGAAG CAGTAGTACA GAAACAGGTG TCGATGATGC
TTACCCAATC
661 CAAGTTAAAC TTGGATCTGG CCTTAGCTTT GACAGTACAG GAGCCATAAT
GGCTGGTAAC
721 AAAGAAGACG ATAACTCAC TTTGTGGACA ACACCTGATC CATCACCAAA
CTGTCAAATA
781 CTCGCAGAAA ATGATGCAAA ACTAACAATT TGCTTGACTA AATGTGGTAG
TCAAATACTG
841 GCCACTGTGT CAGTCTTAGT TGTAGGAAGT GGAGACCTAA ACCCCATTAC
TGGCACCGTA
901 AGCAGTGCTC AGGTGTTTCT ACGTTTTGAT GCAAACGGTG TTCTTTTAAAC
AGAACATTCT
961 AACTAAAAA AATACTGGGG GTATAGGCAG GGAGATAGCA TAGATGGCAC
TCCATATGCC
1021 AATGCTGTAG GATTCATGCC CAATTTAAAA GCTTATCAA AGTCACAAAG
TTCTACTACT
1081 AAAAATAATA TAGTAGGGCA AGTATACATG AATGGAGATG TTTCAAACC
TATGCTTCTC
1141 ACTATAACCC TCAATGGTAC TGATGACAGC AACAGTACAT ATTCAATGTC
ATTTTCATAC

Fig. 14A

ITR0048PV

SEQ ID NO: 10

90/153

1201 ACCTGGACTA ATGGAAGCTA TGTTGGAGCA ACATTGGAG CTA ACTCTTA
TACCTTCTCC

1261 TACATCGCCC AAGAATGA

Fig. 14B

ITR0048PV

SEQ ID NO: 11

91/153

1 ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC
CTACGATGCA

61 GACAACGCAC CGACCGTGCC CTTTCATCAAC CCCCCCTTCG TCTCTTCAGA
TGGATTCCAA

121 GAGAAGCCCC TGGGGGTGCT GTCCCTGCGA CTGGCCGACC CCGTCACCAC
CAAGAACGGG

181 GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA
ACTCATCTCC

241 AACACGGCCA CCAAGGCCGC CGCCCCCTCTC AGTTTTTCCA ACAACACCAT
TTCCCTTAAC

301 ATGGATCACC CCTTTTACAC TAAAGATGGA AAATTATCCT TACAAGTTTC
TCCACCATTA

361 AATATACTGA GAACAAGCAT TCTAAACACA CTAGCTTTAG GTTTTGGATC
AGGTTTAGGA

421 CTCCGTGGCT CTGCCTTGGC AGTACAGTTA GTCTCTCCAC TTACATTGA
TACTGATGGA

481 AACATAAAGC TTACCTTAGA CAGAGGTTTG CATGTTACAA CAGGAGATGC
AATTGAAAGC

541 AACATAAGCT GGGCTAAAGG TTTAAAATTT GAAGATGGAG CCATAGCAAC
CAACATTGGA

601 AATGGGTTAG AGTTTGGAAG CAGTAGTACA GAAACAGGTG TTGATGATGC
TTACCCAATC

661 CAAGTTAAAC TTGGATCTGG CCTTAGCTTT GACAGTACAG GAGCCATAAT
GGCTGGTAAC

721 AAAGAAGACG ATAAACTCAC TTTGTGGACA ACACCTGATC CATCGCCAAA
CTGTCAAATA

781 CTCGCAGAAA ATGATGCAAA ACTAACACTT TGCTTGACTA AATGTGGTAG
TCAAATACTG

841 GCCACTGTGT CAGTCTTAGT TGTAGGAAGT GGAAACCTAA ACCCCATTAC
TGGCACCGTA

901 AGCAGTGCTC AGGTGTTTCT ACGTTTTGAT GCAAACGGTG TTCTTTTAAC
AGAACATTCT

961 AACTAATAAA AATACTGGGG GTATAGGCAG GGAGATAGCA TAGATGGCAC
TCCATATACC

1021 AATGCTGTAG GATTCATGCC CAATTTAAAA GCTTATCCA AGTCACAAAG
TTCTACTACT

1081 AAAAATAATA TAGTAGGGCA AGTATACATG AATGGAGATG TTTCAAACC
TATGCTTCTC

Fig. 15A

ITR0048PV

SEQ ID NO: 11

92/153

1141 ACTATAACCC TCAATGGTAC TGATGACAGC AACAGTACAT ATTCAATGTC
ATTTTCATAC

1201 ACCTGGACTA ATGGAAGCTA TGTGGAGCA ACATTGGGG CTAACCTTA
TACCTTCTCA

1261 TACATCGCCC AAGAATGA

Fig. 15B

ITR0048PV

SEQ ID NO: 12

93/153

1 ATGAAGCGCA CCAAACGTC TGACGAGAGC TTCAACCCCG TGTACCCCTA
TGACACGGAA

61 AACGGTCCTC CCTCCGTCCC TTTCCTCACC CCTCCCTTCG TGTCTCCCGA
TGGATTCCAA

121 GAGAGCCCCC CCGGGGTCCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC
TCCCCACGGC

181 ATGCTCGCCC TGAAATGGG AAGTGGCCTC TCCCTGGACG ACGCCGGCAA
CCTCACCTCT

241 CAAGATGTCA CCACCACTAC CCCTCCCCTG AAAAAACCA AGACCAACCT
CAGCCTAGAA

301 ACCTCAGCCC CCCTGACTGT GAGCACCTCA GGCGCCCTCA CCCTAGCAGC
CGCCGTTCCC

361 CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGAC
AGTCCAAGAT

421 GCAAACTCA CCCTGGCCAC CAAGGGCCCC CTGACCGTGT CTGAAGGCAA
ACTAGCCTTG

481 CAGACCTCGG CCCCCTGAC GGCCGCTGAC AGCAGCACCC TCACAATCAG
CGCCACACCG

541 CCCCTTAGCA CAAGCAATGG CAGCTTGGGT ATTGACATGC AAGCCCCCAT
TTACACTACT

601 AACGGAAAAC TGGGACTTAA CTTTGGTGCT CCCCTGCATG TGGTAGACAG
CCTAAATGCA

661 CTGACTGTAG TGA CTGGCCA AGGTCTTACG ATAAACGGTA CAGCCCTACA
AACTAGAGTC

721 TCAGGTGCCC TCAACTATGA CTCATCAGGA AACCTAGAAT TGAGAGCTGC
AGGGGGTATG

781 CGAGTTGATG CAAATGGCAA ACTTATCCTT GACGTAGCTT ACCCATTTGA
TGCTCAAAAC

841 AACCTCAGCC TTAGACTTGG ACAGGGACCC CTGTTTGTTA ACTCTGCCCCA
CAACTTGGAT

901 GTTAACTACA ACAGAGGCCT CTACCTGTTC ACATCTGGAA ATACCAAAAA
GCTAGAAGTT

961 AATATCAAAA CAGCCAAAGG CCTCATTTAT GATGACACTG CTATAGCAAT
CAATCCAGGC

1021 GATGGGCTAG AGTTTGGCTC AGGCTCAGAT ACAAATCCAT TAAAACTAA
ACTTGGATTG

1081 GGA CTAGAGT ATGACTCCAG CAGAGCCATA ATTGCTAAGC TGGGAACCGG
CCTAAGCTTT

Fig. 16A

ITR0048PV

SEQ ID NO: 12

94/153

1141 GACAACACAG GTGCCATCAC AGTGGGCAAC AAAAATGATG ACAAGCTTAC
CTTGTGGACC

1201 ACACCAGACC CCTCTCCCAA CTGTAGAATT TATTCAGAAA AAGATGCTAA
ATTTACACTA

1261 GTTTTAACTA AATGCGGCAG TCAGGTGTTG GCCAGCGTTT CTGTTTTATC
TGTAAGG

1321 AGCCTTGCGC CCATCAGTGG CACAGTAACT AGCGCTCAGA TTATTCTCAG
ATTTGATGAA

1381 AATGGAGTTC TACTAAGCAA TTCTTCTCTT GACCCCCAAT ACTGGAAC TA
CAGAAAAGGT

1441 GACCTTACAG AGGGCACTGC ATATACCAAC GCAGTGGGAT TTATGCCCAA
CCTCACAGCA

1501 TACCCAAAAA CACAGAGTCA AACTGCTAAA AGCAACATTG TAAGCCAGGT
TTACTTGAAT

1561 GGGGACAAAT CCAAACCCAT GATCCTCACC ATTACCCTCA ATGGAAC TAA
TGAAACAGGG

1621 GATGCTACAG TTAGCACTTA CTCCATGTCA TTCTCATGGA ATTGGAATGG
AAGTAATTAC

1681 ATTAATGAAA CGTTCCAAAC CAACTCTTTC ACCTTCTCCT ACATCGCCCA
AGAATAA

Fig. 16B

ITR0048PV

SEQ ID NO: 13

95/153

1 ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC
CTACGATGCA

61 GACAACGCAC CGACCGTGCC CTTTCATCAAC CCCCCCTTCG TCTCTTCAGA
TGGATTCCAA

121 GAGAAGCCCC TGGGGGTGCT GTCCCTGCGA CTGGCTGACC CCGTCACCAC
CAAGAACGGG

181 GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA
ACTCATCTCC

241 AACACGGCCA CCAAGGCCGC CGCCCCCTCTC AGTTTTTCCA ACAACACCAT
TTCCCTTAAC

301 ATGGATACCC CTTTTTACAC CAAAGATGGA AAATTAACCA TGCAGGTCAC
TGCACCACTA

361 AAGTTAGCAA ACACAGCCAT ATTGAACACA CTAGCTATGG CATATGGAAA
TGGATTAGGT

421 CTAAGCAACA ACGCTCTTAC CGTTCAGTTA CAATCTCCAC TCACCTTTAA
CAACAGCAAG

481 GTTGCAATCA ACCTGGGAAA TGGACCACTA AATGTAACAT CAAACAGACT
TAGCATTAAT

541 TGCAAGAGGG GTGTCTATGT CACCACCACA GGAGATGCAA TTGAAACCAA
CATAAGTTGG

601 TCAAATGCTA TTAAATTTAT AGGAAATGCC ATGGGTGTCA ACATTGATAC
AAACAAAGGC

661 TTGCAATTTG GCACCACCAG CACTGTCACA GATGTGACCA ATGCTTTCCC
CATACAAGTC

721 AAACCTGGGG CTGGTCTTGC ATTTGATAGC ACTGGAGCTA TTGTTGCATG
GAACAAAGAG

781 GATGACAGTC TCACTTTGTG GACTACACCA GATCCATCTC CAAATTGCAA
GATAGCATCT

841 GACAAAGATG CTAAACTCAC ACTTTGCTTG ACAAATGTG GTAGTCAGAT
ACTGGGCACT

901 GTCTCCTTGT TAGCTGTGAG TGGCAGTTA GCTCCTATCA CTGGAGCTGT
GAGCACTGCA

961 CTTGTATCAC TTAAATTCGA TGCCAATGGA GCACTCTTGG AAAAATCAAC
CCTAAACAGA

1021 GAATATTGGA ACTATAGACA AGGAGATCTT ATTCCAGGTA CGCCATATAC
TCACGCAGTA

1081 GGTTCATGC CCAACAAGAA AGCCTACCCCT AAAACACAA CTGCAGCTTC
CAAAGCCAC

Fig. 17A

ITR0048PV

SEQ ID NO: 13

96/153

1141 ATTGTGGGAG AAGTCTATCT AGACGGAGAT GCAGATAAGC CCCTATCTCT
CATAATCACT

1201 TTTAATGAAA CTGATGATGA ATCATGTGAC TATTGCATGA ACTTTCAATG
GAAATGGGGT

1261 GCTGATCAAT ACAAGGACAA AACACTCGCT ACCAGCTCCT TCACCTTCTC
CTACATTGCC

1321 CAAGAATGA

Fig. 17B

ITR0048PV

SEQ ID NO: 14

97/153

1 ATGAAGCGCA CCAAAACGTC TGACGAGAGC TTCAACCCCG TGTACCCCTA
TGACACGGAA

61 AGCGGCCCTC CCTCCGTCCC TTTCTCACC CCTCCCTTCG TGTCTCCCGA
TGGATTCCAA

121 GAAAGCCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC
TTCCCACGGC

181 ATGCTTGCCC TGAAAATGGG AAGTGGCCTC TCCCTGGACG ACGCTGGCAA
CCTTACCTCT

241 CAAGATATTA CCTCCACTAC CCCTCCCCTC AAAAAACCA AGACCAACCT
CAGCCTAGAA

301 ACCTCATCCC CCCTAACTGT AAGCACCTCA GCGCCCTCA CCGTAGCAGC
CGCCGCTCCC

361 CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGGC
AGTACAGGAT

421 GCAAAACTCA CCCTGGCCAC CAAAGGCCCC CTGACCGTGT CTGAAGGCAA
ACTGGCCTTG

481 CAAACATCGG CCCCCTGAC GGCCGCTGAC AGCAGCACCC TCACCGTTAG
CTCCACTCCA

541 CCAATTAGTG TAAGCAGTGG AAGTTTGGGC TTGGACATGG AAGACCCCAT
GTATACTCAC

601 GATGGAAAAC TGGGAATAAG AATTGGGGGT CCACTAAGAG TAGTAGACAG
CTTGCACACA

661 CTCACTGTAG TTACCGGAAA TGGACTAACT GTAGATAACA ATGCCCTCCA
AACTAGAGTT

721 ACGGGCGCCC TAGGTTATGA CACATCAGGA AATCTACAAC TGAGAGCCGC
AGGGGGTATG

781 CGAATTGATG CAAATGGCCA ACTTATCCTT GATGTGGCAT ACCCATTTGA
TGCTCAAAAC

841 AATCTCAGCC TTAGACTTGG TCAGGGACCC CTGTATGTAA ATACAGACCA
CAACCTGGAT

901 TTAAATTGCA ACAGAGGTCT AACCACAACT ACCACCAACA ACACAAAAA
ACTTGAGACT

961 AAAATTAGCT CAGGCTTAGA CTATGACACC AATGGTGCTG TCATTATTAA
ACTTGGCACT

1021 GGTCTAAGCT TCGACAACAC AGGCGCCCTA ACTGTGGGAA AACTGGTGA
TGATAAACTG

1081 ACTCTGTGGA CGACCCAGA CCCATCTCCA AATTGCAGAA TTTACTCAGA
CAAAGACTGC

Fig. 18A

ITR0048PV

SEQ ID NO: 14

98/153

1141 AAGTTTACTC TCGTCCTAAC TAAGTGTGGA AGCCAAATCC TGGCCTCTGT
CGCCGCCCTA

1201 GCGGTATCAG GAAATCTGGC TTCGATAACA GGCACCGTTG CCAGCGTTAC
CATCTTTCTT

1261 AGATTTGATC AGAATGGAGT GCTTATGGAA AACTCCTCAC TAGACAAGCA
GTACTGGAAC

1321 TTCAGAAATG GCAATTCAAC TAATGCTGCC CCCTACACCA ACGCAGTTGG
GTTTCATGCCA

1381 AACCTCGCAG CGTACCCCAA AACGCAAAGC CAGACTGCTA AAAACAACAT
TGTAAGTCAG

1441 GTTTACTTGA ATGGAGACAA ATCCAAACCC ATGACCCTTA CCATCACCTT
CAATGGAACT

1501 AATGAATCCA GTGAAACTAG TCAGGTGAGT CACTACTCCA TGTCATTTAC
ATGGGCTTGG

1561 GAAAGCGGGC AATATGCCAC TGAAACCTTT GCCACCAACT CCTTCACCTT
TTCTTACATT

1621 GCTGAACAAT AA

Fig. 18B

ITR0048PV

SEQ ID NO: 15

99/153

1 ATGAAGCGCA CAAAACGTC TGACAAGAGC TTCAACCCCG TGTACCCCTA
TGACACGGAA

61 AACGGTCCTC CCTCCGTCCC TTTCTCACC CCTCCCTTCG TGTCTCCCGA
TGGATTCCAA

121 GAGAGCCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC
TTCCCACGGC

181 ATGCTCGCCC TGAAATGGG AAGTGGCCTC TCCCTGGACG ACGCCGGCAA
CCTCACCTCT

241 CAAGATGTCA CCACCACTAC CCCTCCCCTG AAAAAACCA AGACCAACCT
CAGCCTAGAA

301 ACCTCAGCCC CCCTGACTGT GAGCACCTCA GCGGCCCTCA CCCTAGCAGC
CGCCGCCCCC

361 CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGAC
AGTCCAAGAT

421 GCAAACTCA CCCTGGCCAC CAAGGGCCCC CTGACCGTGT CTGAAGGCAA
ACTGGCCTTG

481 CAGACCTCGG CCCCCTGAC GGCCGCTGAC AGCAGCACCC TCACCGTTAG
CGCCACACCA

541 CCCATCAGTG TAAGCAGTGG AAGTTTGGGC TTAGACATGG AAGACCCAAT
GTATACTCAT

601 GATGGAAAAC TGGGAATAAG AATTGGGGGC CCACTGAGAG TAGTAGACAG
CCTGCACACA

661 CTGACTGTAG TTACCGGAAA TGAATAGCT GTAGATAACA ATGCCCTCCA
AACTAGAGTT

721 ACGGGCGCCC TGGGTATGA CACATCAGGA AACCTACAAC TGAGAGCCGC
GGGGGGTATG

781 CGAATTGATG CAAATGGCCA ACTTATCCTT GATGTGGCAT ACCCATTTGA
TGCTCAAAAC

841 AATCTCAGCC TTAGACTTGG TCAGGGACCC CTGTATGTAA ACACAGACCA
CAACCTAGAT

901 TTGAATTGCA ACAGAGGTCT GACCACAACCT ACCACCAACA ACACAAAAAA
ACTTGAAACT

961 AAAATTGGCT CAGGCTTAGA CTATGATACC AATGGTGCTG TTATTATTAA
ACTTGGCACT

1021 GGTGTCAGCT TTGACAGCAC AGGTGCCCTA AGTGTGGGAA AACTGGCGA
TGATAAACTG

1081 ACTCTGTGGA CAACCCAGA CCCATCTCCA AATGCAGAA TTCACTCAGA
CAAAGACTGC

Fig. 19A

ITR0048PV

SEQ ID NO: 15

100/153

1141 AAGTTTACTC TAGTCCTAAC TAAGTGTGGA AGTCAAATCC TGGCTTCTGT
CGCCGCCCTA

1201 GCGGTGTCAG GAAATCTGGC TTCAATAACA GGCACCGTTT CCAGCGTTAC
CATCTTTCTC

1261 AGATTTGATC AGAATGGAGT GCTTATGGAA AACTCCTCGC TAGACAAGCA
GTACTGGAAC

1321 TTCAGAAATG GTAATTCAAC CAATGCCACC CCCTACACCA ATGCAGTTGG
GTTTATGCCA

1381 AACCTCGCAG CATACCCCAA GACACAGAGC CAGACTGCAA AAAACAACAT
TGTAAGTCAG

1441 GTTTACTTGA ATGGGGACAA ATCCAAACCC ATGACCCTTA CCATTACCCT
CAATGGAAC

1501 AATGAATCCA GTGAACTAG CCAGGTGAGT CACTACTCCA TGTCATTAC
GTGGGCTTGG

1561 GAGAGTGGGC AATATGCCAC CGAAACCTTT GCCACCAATT CCTTTACCTT
CTCTTACATT

1621 GCTGAACAAT AA

Fig. 19B

1	MAKTRILSSS	.FNPVYPYED	EN..SSHPEI	NPGEISPNGF	TQSPDGVLT	NCVAPLTAN	GALDIKVGSG	LKVNSTDGFL	EENIN.....	100
C1	ChAd11	MRTKTSDS	.FNPVYPYDT	ESGPPSVPEL	TPPEVSPDGF	QESPPGVLSL	LAPELVTS	LAPELVTS	LAPELVTS
ChAd20	MRTKTSDS	.FNPVYPYDT	ESGPPSVPEL	TPPEVSPDGF	QESPPGVLSL	LAPELVTS	LAPELVTS	LAPELVTS	LAPELVTS
ChAd17	MRTKTSDS	.FNPVYPYDT	ESGPPSVPEL	TPPEVSPDGF	QESPPGVLSL	LAPELVTS	LAPELVTS	LAPELVTS	LAPELVTS
ChAd3	MRTKTSDS	.FNPVYPYDT	ESGPPSVPEL	TPPEVSPDGF	QESPPGVLSL	LAPELVTS	LAPELVTS	LAPELVTS	LAPELVTS
ChAd19	MRTKTSDS	.FNPVYPYDT	ESGPPSVPEL	TPPEVSPDGF	QESPPGVLSL	LAPELVTS	LAPELVTS	LAPELVTS	LAPELVTS
PAN6	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd5	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd6	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd7	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
PAN5	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
PAN7	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd9	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd10	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd4	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
CV68	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
ChAd16	MSKKRVRVDD	DEDPVYPYDA	DN.APTVPEI	NPPEVSSDGF	QEKPLGVLSL	RLADPVTTKN	GEITLKGEG	VDLDSGGKLI	SNTAT.....
101	200
C1	ETSAPLTVST	SGALTAAAV	PLAVAGTSIT	MQSEAPLTVQ	DAKLTATKG	PLTVSEKGLA	LOTSAPLTA	DSSTLTISAT	PPLSTSNGL	GIDMQAPIYT
ChAd11	ETSAPLTVST	SGALTAAAV	PLAVAGTSIT	MQSEAPLTVQ	DAKLTATKG	PLTVSEKGLA	LOTSAPLTA	DSSTLTISAT	PPLSTSNGL	GIDMQAPIYT
ChAd20	ETSAPLTVST	SGALTAAAV	PLAVAGTSIT	MQSEAPLTVQ	DAKLTATKG	PLTVSEKGLA	LOTSAPLTA	DSSTLTISAT	PPLSTSNGL	GIDMQAPIYT
ChAd17	ETSAPLTVST	SGALTAAAV	PLAVAGTSIT	MQSEAPLTVQ	DAKLTATKG	PLTVSEKGLA	LOTSAPLTA	DSSTLTISAT	PPLSTSNGL	GIDMQAPIYT
ChAd3	ETSAPLTVST	SGALTAAAV	PLAVAGTSIT	MQSEAPLTVQ	DAKLTATKG	PLTVSEKGLA	LOTSAPLTA	DSSTLTISAT	PPLSTSNGL	GIDMQAPIYT
ChAd19	ETSAPLTVST	SGALTAAAV	PLAVAGTSIT	MQSEAPLTVQ	DAKLTATKG	PLTVSEKGLA	LOTSAPLTA	DSSTLTISAT	PPLSTSNGL	GIDMQAPIYT
PAN6
ChAd5
ChAd6
ChAd7
PAN5
PAN7
ChAd9
ChAd10
ChAd4
CV68
ChAd16

Fig. 20A

201	300
C1
Chad11	TNGKLGNGF APLHVVD..
Chad20	TNGKLGNGF APLHVVD..
Chad17	HDGKLGIRIG GPLRVVD..
Chad3	HDGKLGIRIG GPLRVVD..
Chad19	HDGKLGIRIG GPLRVVD..
PAN6	HDGKLGIRIG GPLRVVD..
PAN6	HDGKLGIRIG GPLRVVD..
Chad5	HDGKLGIRIG GPLRVVD..
Chad6	HDGKLGIRIG GPLRVVD..
Chad7	HDGKLGIRIG GPLRVVD..
PAN5	HDGKLGIRIG GPLRVVD..
PAN7	HDGKLGIRIG GPLRVVD..
Chad9	HDGKLGIRIG GPLRVVD..
Chad10	HDGKLGIRIG GPLRVVD..
Chad4	HDGKLGIRIG GPLRVVD..
CV68	HDGKLGIRIG GPLRVVD..
Chad16	HDGKLGIRIG GPLRVVD..
301	400
C1
Chad11	FVNSAHNLDV NYNRGLYFT
Chad20	FVNSAHNLDV NYNRGLYFT
Chad17	YVNTDHNLDL NCNRGLYFT
Chad3	YVNTDHNLDL NCNRGLYFT
Chad19	YVNTDHNLDL NCNRGLYFT
PAN6	YVNTDHNLDL NCNRGLYFT
Chad5	YVNTDHNLDL NCNRGLYFT
Chad6	YVNTDHNLDL NCNRGLYFT
Chad7	YVNTDHNLDL NCNRGLYFT
PAN5	YVNTDHNLDL NCNRGLYFT
PAN7	YVNTDHNLDL NCNRGLYFT
Chad9	YVNTDHNLDL NCNRGLYFT
Chad10	YVNTDHNLDL NCNRGLYFT
Chad4	YVNTDHNLDL NCNRGLYFT
CV68	YVNTDHNLDL NCNRGLYFT
Chad16	YVNTDHNLDL NCNRGLYFT

Fig. 20B

401
 C1NLTWG AKPSANCVIK EGEDSPDCKL TIVLVKNGGL VNGYITIMGD SE.YTNTLFK NKQVTTIDVNL AFNTGQIIT YLSLKS.NL NFKDNQWAT
 ChAd11 NDDKLTWT PDPSPNCRIY SE...KDAKE TLVLTCKGSQ VLASVSLV K..GSLAPIS GTVTSQIIVL RFDENGVLIS N.SSLDPQYW NYRK.GDLTE
 ChAd20 NDDKLTWT PDPSPNCRIY SE...KDAKE TLVLTCKGSQ VLASVSLV K..GSLAPIS GTVTSQIIVL RFDENGVLIS N.SSLDPQYW NYRK.GDLTE
 ChAd17 GDDKLTWT PDPSPNCRIH SD...KDCKE TLVLTCKGSQ ILASVAALAV S..GNLASIT GTVASVTIFL RFDQNGVIME N.SSLDRQYW NFRN.GNSTN
 ChAd3 GDDKLTWT PDPSPNCRIH SD...KDCKE TLVLTCKGSQ ILASVAALAV S..GNLASIT GTVASVTIFL RFDQNGVIME N.SSLDRQYW NFRN.GNSTN
 ChAd19 GDDKLTWT PDPSPNCRIH SD...KDCKE TLVLTCKGSQ ILASVAALAV S..GNLASIT GTVASVTIFL RFDQNGVIME N.SSLDRQYW NFRN.GNSTN
 PAN6 QDDKLTWT PDPSPNCQLL SD...RDAKE TLCLTRCKGSQ ILGTVAVAL TVGSALPIN DTVKSALVEL RFDSDGVIMS N.SSMVGDIW NERE.GDTTQ
 ChAd5 EDDKLTWT ADPSPNCHY SA...KDAKL TLCLTRCKGSQ ILGTVSLAV S..GSLAPIT GAVTALVSL KFDANGVILQ K.STLNKEYW NYRK.GDLIP
 ChAd6 EDDKLTWT ADPSPNCHY SA...KDAKL TLCLTRCKGSQ ILGTVSLAV S..GSLAPIT GAVTALVSL KFDANGVILQ K.STLNKEYW NYRK.GDLIP
 ChAd7 EDDKLTWT ADPSPNCHY SD...KDAKL TLCLTRCKGSQ ILGTVSLAV S..GSLAPIT GAVTALVSL KFDANGVILQ K.STLNKEYW NYRK.GDLIP
 PAN5 EDDKLTWT ADPSPNCHY SE...KDAKL TLCLTRCKGSQ ILGTVSLAV S..GSLAPIT GAVTALVSL KFDANGVILQ K.STLNKEYW NYRK.GDLIP
 PAN7 EDDKLTWT ADPSPNCKIY SE...KDAKL TLCLTRCKGSQ ILGTVSLAV S..GSLAPIT GAVTALVSL KFDANGVILQ K.STLNKEYW NYRK.GDLIP
 ChAd9 EDDKLTWT PDPSPNCQIL AE...NDAKL TLCLTRCKGSQ ILATVSLV GS.GDLNPIT GTVSSAQVEL RFDANGVILLT E.HSTLKKYW GYRO.GDSID
 ChAd10 EDDKLTWT PDPSPNCQIL AE...NDAKL TLCLTRCKGSQ ILATVSLV GS.GDLNPIT GTVSSAQVEL RFDANGVILLT E.HSTLKKYW GYRO.GDSID
 ChAd4 EDDKLTWT PDPSPNCQIL AE...NDAKL TLCLTRCKGSQ ILATVSLV GS.GDLNPIT GTVSSAQVEL RFDANGVILLT E.HSTLKKYW GYRO.GDSID
 CV68 EDDKLTWT PDPSPNCQIL AE...NDAKL TLCLTRCKGSQ ILATVSLV GS.GDLNPIT GTVSSAQVEL RFDANGVILLT E.HSTLKKYW GYRO.GDSID
 ChAd16 EDDSLTWT PDPSPNCKIA SD...KDAKL TLCLTRCKGSQ ILGTVSLAV S..GSLAPIT GAVTALVSL KFDANGVILQ K.STLNKEYW NYRK.GDLIP
 500
 501
 C1 GTIT.SAKGE MPSTTAYPFI TYATQSLN.E DYIYGEYK STNGTLEPLK VVTILNRMS AS....CMAY AMNFSWSLNA EEPETTEVT LITSPEFFSY
 ChAd11 GTAYTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE TGD.ATVSTY SMSFSWNWG S...NYINET FQTNSTFSY
 ChAd20 GTAYTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE TGD.ATVSTY SMSFSWNWG S...NYINET FQTNSTFSY
 ChAd17 AAPTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 ChAd3 AAPTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 ChAd19 AAPTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 PAN6 SVAYTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 ChAd5 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd6 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd7 AEPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 PAN5 AEPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 PAN7 AEPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd9 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd10 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd4 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 CV68 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd16 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 600
 601
 C1 GTIT.SAKGE MPSTTAYPFI TYATQSLN.E DYIYGEYK STNGTLEPLK VVTILNRMS AS....CMAY AMNFSWSLNA EEPETTEVT LITSPEFFSY
 ChAd11 GTAYTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE TGD.ATVSTY SMSFSWNWG S...NYINET FQTNSTFSY
 ChAd20 GTAYTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE TGD.ATVSTY SMSFSWNWG S...NYINET FQTNSTFSY
 ChAd17 AAPTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 ChAd3 AAPTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 ChAd19 AAPTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 PAN6 SVAYTNAVGE MPNLTAYP... ..KTQSOTAK SNIVSQVILN GD..KSRPMI LITTLNGTNE SSETSQVSHY SMSFTWAWES G...QYATET FATNSFTFSY
 ChAd5 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd6 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd7 AEPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 PAN5 AEPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 PAN7 AEPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd9 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd10 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd4 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 CV68 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY
 ChAd16 GTPYTHAVGE MPNLTAYP... ..KNTAAAK SHIVGKVYLH GD..ESKPLD LIITFNETSD ES.....CTY CINFQWQWGT D...QYKDET LAVSFTFSY

Fig. 20C

601
C1 IREDD
ChAd11 IAQE.
ChAd20 IAQE.
ChAd17 IAEQ.
ChAd3 IAEQ.
ChAd19 IAEQ.
PAN6 IAQE.
ChAd5 IAQE.
ChAd6 IAKE.
ChAd7 IAQE.
PAN5 IAKE.
PAN7 IAQE.
ChAd9 IAQE.
ChAd10 IAQE.
ChAd4 IAQE.
CV68 IAQE.
ChAd16 IAQE.

Fig. 20D

ITR0048PV

SEQ ID NO: 16

105/153

1 ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG
CCAGGACGCC

61 TCGGAGTACC TGAGCCCCGG GCTGGTGCAG TTCGCCCGCG CCACCGAGAG
CTACTTCAGC

121 CTGAGTAACA AGTTTAGGAA CCCACGGTG GCGCCCACGC ACGATGTGAC
CACCGACCGG

181 TCTCAGCGCC TGACGCTGCG GTTCATTCCC GTGGACCGCG AGGACACCGC
GTACTCGTAC

241 AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC
CTCCACCTAC

301 TTTGACATCC GCGGGGTGCT GGACCGGGGT CCCACTTTCA AGCCCTACTC
TGGCACCGCC

361 TACAACTCCC TGGCCCCCAA GGGCGCTCCC AACCCATGCG AGTGGGATGA
GGCTGCTACT

421 GCCCTTGACA TTGATTTGAA CGCAGAAGAC GATGAAGAAA GCGACGAAGC
TCAAGGGGAA

481 GCAGATCAGC AGAAAACTCA TGTATTTGGC CAGGCGCCCT ACTCCGGACA
GAACATTACA

541 AAAGAAGGCA TACAGATAGG CATAGATGCT GCCAGTCAAG CCCAGACACC
TGTATATGCC

601 GATAAAACAT TCCAACCAGA ACCTCAAGTT GGAGAATCAC AGTGGAATGA
GACAGAGATT

661 AGTTATGGAG CGGGACGGGT GCTTAAAAA ACCACTCTCA TGAAACCTTG
CTATGGGTCG

721 TATGCAAGGC CTACTAATGA GAACGGAGGT CAGGGCATCC TCTTGGAACA
AGATGGAAAG

781 AAAGAAAGTC AAGTGGAAT GCAATTTTTC TCTACTACTC AGGCAGCCGC
GGGTAATTCA

841 GATAATCCTA CCCCAAAGGT TGTTTTGTAC AGCGAGGATG TTAACCTGGA
AACACCAGAT

901 ACACACATTT CATAATGCC CACCAACAAC GAGACAAATT CAAGAGAGCT
TTTGGGACAA

961 CAGGCCATGC CCAACAGGCC TAATTACATT GGCTTCAGAG ACAACTTTAT
CGGTCTCATG

1021 TATTACAACA GCACTGGCAA CATGGGAGTG CTTGCAGGTC AGGCCTCTCA
GTTGAACGCA

1081 GTGGTGGACT TGCAAGACAG AAACACAGAA CTGTCATACC AGCTCTTGCT
TGATTCCATG

Fig. 21A

ITR0048PV

SEQ ID NO: 16

106/153

1141 GGTGACAGAA CCAGATACTT TTCCATGTGG AATCAGGCAG TGGACAGTTA
TGACCCAGAT

1201 GTCAGAATTA TTGAAAATCA TGGAAGTGA GACGAGCTCC CCAACTATTG
TTTCCCTCTG

1261 GGCGGCGTAA TCAATACGGA AACTTTCACA AAAGTAAAC CTAAAGCTGC
ACAGGACGCT

1321 CAGTGGGAAA AAGATTCAGA ATTTTCAGAT AAAAATGAAA TAAGGGTGGG
AAACAACCTC

1381 GCCATGGAAA TTAACCTCAA TGCCAATCTG TGGAGGAACT TTTTGTACTC
CAACGTAGCC

1441 CTCTACTTGC CTGACAAGCT TAAGTATACT CCATCCAATG TGCAAATTTT
CAACAATCCC

1501 AACTCCTACG ATTACATGAA CAAGCGAGTG GTGGCCCCGG GGCTGGTGGA
CTGCTACATC

1561 AACCTGGGCG CGCGCTGGTC GCTGGACTAC ATGGACAACG TCAACCCCTT
CAACCACCAC

1621 CGCAATGCGG GCCTGCGCTA CCGCTCCATG CTCCTGGGCA ACGGGCGCTA
CGTGCCCTTC

1681 CACATCCAGG TGCCCCAGAA GTTCTTTGCC ATCAAGAACC TCCTCCTCCT
GCCGGGCTCC

1741 TACACCTACG AGTGGAACCT CAGGAAGGAT GTCAACATGG TCCTCCAGAG
CTCTCTGGGT

1801 AACGATCTCA GGGTGGACGG GGCCAGCATC AAGTTCGAGA GCATCTGCCT
CTACGCCACC

1861 TTCTTCCCCA TGGCCCACAA CACGGCCTCC ACGCTCGAGG CCATGCTCAG
GAACGACACC

1921 AACGACCAGT CCTTCAATGA CTACCTTTC GCCGCAACA TGCTCTACCC
CATACCCGCC

1981 AACGCCACCA ACGTCCCCAT CTCCATCCCC TCGCGCAAAT GGGCGGCCTT
CCGCGGCTGG

2041 GCCTTCACCC GCCTCAAGAC CAAGGAGACC CCCTCCCTGG GCTCGGGATT
CGACCCCTAC

2101 TACACCTACT CGGGCTCCAT TCCCTACCTG GACGGCACCT TCTACCTCAA
CCACACTTTC

2161 AAGAAGGTCT CGGTCACCTT CGACTCCTCG GTCAGCTGGC CGGGCAACGA
CCGTCTGCTC

Fig. 21B

ITR0048PV

SEQ ID NO: 16

107/153

2221 ACCCCCAACG AGTTCGAGAT CAAGCGCTCG GTCGACGGGG AGGGCTACAA
CGTGGCCCAG

2281 TGCAACATGA CCAAGGACTG GTTCCTGGTC CAGATGCTGG CCAACTACAA
CATCGGCTAC

2341 CAGGGCTTCT ACATCCCAGA GAGCTACAAG GACAGGATGT ACTCCTTCTT
CAGGAAC TTC

2401 CAGCCCATGA GCCGGCAGGT GGTGGACCAG ACCAAGTACA AGGACTACCA
GGAGGTGGGC

2461 ATCATCCACC AGCACAACAA CTCGGGCTTC GTGGGCTACC TCGCCCCAC
CATGCGCGAG

2521 GGACAGGCCT ACCCCGCCAA CTTCCCCTAC CCGCTCATAG GCAAGACCGC
GGTCGACAGC

2581 ATCACCACAGA AAAAGTTCCT CTGCGACCGC ACCCTCTGGC GCATCCCCTT
CTCCAGCAAC

2641 TTCATGTCCA TGGGTGCGCT CTCGGACCTG GGCCAGAACT TGCTCTACGC
CAACTCCGCC

2701 CACGCCCTCG ACATGACCTT CGAGGTCGAC CCCATGGACG AGCCCACCCT
TCTCTATGTT

2761 CTGTTCAAG TCTTTGACGT GGTCCGGGTC CACCAGCCGC ACCGCGGCGT
CATCGAGACC

2821 GTGTACCTGC GTACGCCCTT CTCGGCCGGC AACGCCACCA CCTAA

Fig. 21C

ITR0048PV

SEQ ID NO: 17

108/153

1 ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG
ACAGGACGCT

61 TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCC GCG CCACAGACAC
CTACTTCAGT

121 CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCACGC ACGATGTGAC
CACCGACCGC

181 AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC
CTACTCGTAC

241 AAAGTGC GCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC
CAGCACCTAC

301 TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCTACTC
CGGCACCGCC

361 TACAACAGCC TGGCTCCCAA GGGAGCGCCC AATTCCAGCC AGTGGGAGCA
AAAAAAGACT

421 GGCAATAATG CCAATGGAGA TACGGAGAAT GTCAC TTATG GTGTAGCTGC
CATGGGAGGA

481 ATTGACATCG ATAAAAATGG CCTTCAAATT GGAACCGATG ACACCAAAGA
TGACGATAAT

541 GAAATTTATG CAGACAAAAC ATATCAGCCT GAGCCGCAA TAGGAGAGGA
AAACTGGCAA

601 GAAACATATT CCTACTATGG AGGTAGAGCT CTTAAAAAAG ATACCAAAT
GAAGCCATGC

661 TATGGCTCAT TTGCCAGACC TACCAATGTG AAAGGAGGAC AGGCAAAAAT
AAAAACAGAT

721 GGAGATGTTA AGTCATTTGA CATAGACCTA GCCTTCTTTG ATATTCCCAA
TTCTGGCGCG

781 GGAAATGGCA CAAATGTTAA CGATGATCCA GATATGGTTA TGTATACAGA
AAATGTAAAT

841 CTGGAAACCC CAGATACTCA TATTGTGTAC AAACCAGGAA CTTCAGATGA
CAGCTCAAAG

901 GTCAACTTGT GTCAGCAATC CATGCCTAAC AGACCCAATT ATATTGGCTT
CAGAGACAAT

961 TTTATTGGGC TTATGTACTA CAACAGCACT GGCAATATGG GTGTGCTGGC
TGGTCAGGCC

1021 TCTCAACTGA ATGCCGTGGT GGACTTGCAA GACAGAAACA CAGAGCTGTC
CTACCAGCTC

1081 TTGCTTGACT CTCTGGGTGA CAGAACCAGG TATTTAGTA TGTGGAATCA
GGCGGTGGAC

Fig. 22A

ITR0048PV

SEQ ID NO: 17

109/153

1141 AGTTATGATC CTGATGTGCG CATTATTGAA AACCATGGTG TGGAGGATGA
ATTGCCAAAC

1201 TATTGCTTCC CTTGGATGG AGCAGGCACC AATTCGGTTT ACCAAGGTGT
TAAACCAAAA

1261 ACTGACAATG GCAACGATCA GTGGGAAACA GATTCCACAG TTTCAAGTCA
CAATCAGATA

1321 TGCAAAGGCA ATATCTATGC CATGGAGATC AACCTCCAGG CCAACCTGTG
GAGAAGTTTT

1381 CTCTACTCGA ACGTGGCCCT GTACCTGCCC GATTCTTACA AGTACACGCC
GGCCAACATC

1441 ACCCTGCCCC CCAACACCAA CACCTACGAT TACATGAACG GGAGAGTGGT
GCCTCCCTCG

1501 CTGGTGGACG CTTACATCAA CATCGGGGCG CGCTGGTCGC TGGACCCCAT
GGACAACGTG

1561 AATCCCTTCA ACCACCACCG CAACGCGGGC CTGCGCTACC GCTCCATGCT
CCTGGGCAAC

1621 GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAAT TTTTCGCCAT
CAAGAGCCTC

1681 CTGCTCCTGC CCGGGTCCTA CACCTACGAG TGGAACCTCC GCAAGGACGT
CAACATGATC

1741 CTGCAGAGCT CCCTCGGCAA CGACCTGCGC ACGGACGGGG CCTCCATCTC
CTTCAACCAGC

1801 ATCAACCTCT ACGCCACCTT CTTCCCCATG GCGCACAACA CGGCCTCCAC
GCTCGAGGCC

1861 ATGCTGCGCA ACGACACCAA CGACCAGTCC TTCAACGACT ACCTCTCGGC
GGCCAACATG

1921 CTCTACCCCA TCCCGGCCAA CGCCACCAAC GTGCCCATCT CCATCCCCTC
GCGCAACTGG

1981 GCCGCCTTCC GCGGCTGGTC CTTACGCGC CTCAAGACCC GCGAGACGCC
CTCGCTGGGC

2041 TCCGGGTTCG ACCCCTACTT CGTCTACTCG GGCTCCATCC CCTACCTCGA
CGGCACCTTC

2101 TACCTCAACC ACACCTTCAA GAAGGTCTCC ATCACCTTCG ACTCCTCCGT
CAGCTGGCCC

2161 GGCAACGACC GCCTCCTGAC GCCCAACGAG TTCGAAATCA AGCGCACCGT
CGACGGAGAG

2221 GGATACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA
GATGCTGGCC

Fig. 22B

ITR0048PV

SEQ ID NO: 17

110/153

2281 CACTACAACA TCGGCTACCA GGGCTTCTAC GTGCCCAGAGG GCTACAAGGA
CCGCATGTAC

2341 TCCTTCTTCC GCAACTTCCA GCCCATGAGC CGCCAGGTGG TGGACGAGGT
CAACTACAAG

2401 GACTACCAGG CCGTCACCCT GGCCTACCAG CACAACAAC TCGGGCTTCGT
CGGCTACCTC

2461 GCGCCCACCA TGCGCCAGGG CCAGCCCTAC CCCGCCAACT ACCCGTACCC
GCTCATCGGA

2521 AAGAGCGCCG TCACCAGCGT CACCCAGAAA AAGTTCCTCT GCGACAGGGT
CATGTGGCGC

2581 ATCCCCTTCT CCAGCAACTT CATGTCCATG GGCGCGCTCA CCGACCTCGG
CCAGAACATG

2641 CTCTATGCCA ACTCCGCCCA CGCGCTAGAC ATGAATTTCTG AAGTCGACCC
CATGGATGAG

2701 TCCACCCCTC TCTATGTTGT CTTTGAAGTC TTCGACGTCG TCCGAGTGCA
CCAGCCCCAC

2761 CGCGGCGTCA TCGAGGCCGT CTACCTGCGC ACCCCCTTCT CGGCCGGTAA
CGCCACCACC

2821 TAA

Fig. 22C

ITR0048PV

SEQ ID NO: 18

111/153

1 ATGGCCACCC CATCGATGCT GCCCAGTGG GCGTACATGC ACATCGCCGG
ACAGGACGCT

61 TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCCGG CCACAGACAC
CTACTTCAGT

121 CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCACGC ACGATGTGAC
CACCGACCGC

181 AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC
CTACTCGTAC

241 AAAGTGCCT ACACGCTGGC CGTGGGCGAC AACCGCGTC TGGACATGGC
CAGCACCTAC

301 TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCTACTC
CGGCACCGCC

361 TACAACAGCC TGGCTCCCAA GGGAGCGCCC AATTCAGCC AGTGGGAGCA
AAAAAAGACT

421 GGCAATAATG CCAATGGAGA TACGGAGAAT GTCACTTATG GTGTAGCTGC
CATGGGAGGA

481 ATTGACATCG ATAAAAATGG CTTCAAATT GGAACCGATG ACACCAAAGA
TGACGATAAT

541 GAAATTTATG CAGACAAAAC ATATCAGCCT GAGCCGCAA TAGGAGAGGA
AAACTGGCAA

601 GAAACATATT CCTACTATGG AGGTAGAGCT CTTAAAAAAG ATACCAAAT
GAAGCCATGC

661 TATGGCTCAT TTGCCAGACC TACCAATGTG AAAGGAGGAC AGGCAAAAT
AAAAACAGAT

721 GGAGATGTTA AGTCATTTGA CATAGACCTA GCCTTCTTTG ATATTCCCAA
TTCTGGCGCG

781 GGAAATGGCA CAAATGTAA CGATGATCCA GATATGGTTA TGTATACAGA
AAATGTAAAT

841 CTGGAAACCC CAGATACTCA TATTGTGTAC AAACCAGGAA CTTCAGATGA
CAGCTCAAAG

901 GTCAACTTGT GTCAGCAATC CATGCCTAAC AGACCCAATT ATATTGGCTT
CAGAGACAAT

961 TTTATTGGGC TTATGTACTA CAACAGCACT GGCAATATGG GTGTGCTGGC
TGGTCAGGCC

1021 TCTCAACTGA ATGCCGTGGT GGACTTGCAA GACAGAAACA CAGAGCTGTC
CTACCAGCTC

1081 TTGCTTGA CTCTGGGTGA CAGAACCAGG TATTTAGTA TGTGGAATCA
GGCGGTGGAC

Fig. 23A

ITR0048PV

SEQ ID NO: 18

112/153

1141 AGTTATGATC CTGATGTGCG CATTATTGAA AACCATGGTG TGGAGGATGA
ATTGCCAAAC

1201 TATTGCTTCC CTTGGATGG AGCAGGCACC AATTCGGTTT ACCAAGGTGT
TAAACCAAAA

1261 ACTGACAATG GCAACGATCA GTGGGAAACA GATTCCACAG TTTCAAGTCA
CAATCAGATA

1321 TGCAAAGGCA ATATCTATGC CATGGAGATC AACCTCCAGG CCAACCTGTG
GAGAAGTTTT

1381 CTCTACTCGA ACGTGGCCCT GTACCTGCCC GATTCTTACA AGTACACGCC
GGCCAACATC

1441 ACCCTGCCCC CCAACACCAA CACCTACGAT TACATGAACG GGAGAGTGGT
GCCTCCCTCG

1501 CTGGTGGACG CCTACATCAA CATCGGGGCG CGCTGGTCGC TGGACCCCAT
GGACAACGTG

1561 AATCCCTTCA ACCACCACCG CAACGCGGGC CTGCGCTACC GCTCCATGCT
CCTGGGCAAC

1621 GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAAT TTTTGGCCAT
CAAGAGCCTC

1681 CTGCTCCTGC CCGGGTCCTA CACCTACGAG TGGAACCTCC GCAAGGACGT
CAACATGATC

1741 CTGCAGAGCT CCCTCGGCAA CGACCTGCGC ACGGACGGGG CCTCCATCTC
CTTCACCAGC

1801 ATCAACCTCT ACGCCACCTT CTTCCCCATG GCGCACAACA CGGCCTCCAC
GCTCGAGGCC

1861 ATGCTGCGCA ACGACACCAA CGACCAGTCC TTCAACGACT ACCTCTCGGC
GGCCAACATG

1921 CTCTACCCCA TCCCGGCCAA CGCCACCAAC GTGCCCATCT CCATCCCCTC
GCGCAACTGG

1981 GCCGCCTTCC GCGGCTGGTC CTTACGCGC CTCAAGACCC GCGAGACGCC
CTCGCTGGGC

2041 TCCGGGTTCG ACCCCTACTT CGTCTACTCG GGCTCCATCC CCTACCTCGA
CGGCACCTTC

2101 TACCTCAACC ACACCTTCAA GAAGGTCTCC ATCACCTTCG ACTCCTCCGT
CAGCTGGCCC

2161 GGCAACGACC GCCTCCTGAC GCCCAACGAG TTCGAAATCA AGCGCACCGT
CGACGGAGAG

2221 GGATACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA
GATGCTGGCC

Fig. 23B

ITR0048PV

SEQ ID NO: 18

113/153

2281 CACTACAACA TCGGCTACCA GGGCTTCTAC GTGCCCAGAG GCTACAAGGA
CCGCATGTAC

2341 TCCTTCTTCC GCAACTTCCA GCCCATGAGC CGCCAGGTCG TGGACGAGGT
CAACTACAAG

2401 GACTACCAGG CCGTCACCTT GGCCTACCAG CACAACAACT CGGGCTTCGT
CGGCTACCTC

2461 GCGCCACCA TCGCCAGGG CCAGCCCTAC CCCGCCAACT ACCCCTACCC
GCTCATCGGC

2521 AAGAGCGCCG TCGCCAGCGT CACCCAGAAA AAGTTCCTCT GCGACCGGGT
CATGTGGCGC

2581 ATCCCCTTCT CCAGCAACTT CATGTCCATG GGCGCGCTCA CCGACCTCGG
CCAGAACATG

2641 CTCTACGCCA ACTCCGCCCA CGCGCTAGAC ATGAATTTTCG AAGTCGACCC
CATGGATGAG

2701 TCCACCCTTC TCTATGTTGT CTCGAAGTC TTCGACGTCG TCCGAGTGCA
CCAGCCCCAC

2761 CGCGGCGTCA TCGAGGCCGT CTACCTGCGC ACCCCCTTCT CGGCCGGTAA
AGCCACCACC

2821 TAA

Fig. 23C

ITR0048PV

SEQ ID NO: 19

114/153

1 ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG
ACAGGACGCT

61 TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCC GCG CCACAGACAC
CTACTTCAGT

121 CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC
CACCGACCGC

181 AGCCAGCGGC TGACGCTGCG CTTCTGTCCT GTGGACCGCG AGGACAACAC
CTACTCGTAC

241 AAAGTGCCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC
CAGCACCTAC

301 TTTGACATCC GCGGCGTGCT GGACCGGGC CCTAGCTTCA AACCTTACTC
CGGCACCGCT

361 TACAACAGCC TGGCCCCCAA GGGAGCACCC AATTCAGCC AGTGGGAGCA
AAAAAAGACT

421 GGCAAAAATG CCAATGGAGA TACGGAGAAT GTCACCTATG GTGTAGCTGC
CATGGGAGGA

481 ATTGACATCG ATAAAAATGG CCTTCAAATT GGAACCGATG ACACCAAAGA
TGGCGATAAT

541 GAAATTTATG CAGACAAAAC ATATCAGCCT GAGCCGCAA TAGGAGAGGA
AAACTGGCAA

601 GAAACATATT CCTACTATGG AGGTAGAGCT CTTAAAAAG ATACCAAAT
GAAGCCATGC

661 TATGGCTCAT TTGCTAGACC TACCAATGTG AAAGGAGGAC AGGCAAAAAT
AAAAACAGAT

721 GGAGATGTTA AGTCATTTGA CATAGACCTA GCCTTCTTG ATATTCCAAA
TTCTGGCGCG

781 GGAAATGGCA CAAATGTAA CGATGATCCA GATATGGTTA TGTATACAGA
AAATGTAAAT

841 CTGGAAACCC CAGATACTCA TATTGTGTAC AAACCAGGAA CTTTCTAGTA
CAGCTCCGAG

901 GTCAACTTGT GTCAGCAATC CATGCCTAAC AGACCCAATT ATATTGGCTT
CAGAGACAAT

961 TTTATTGGGC TTATGTACTA CAACAGCACT GGCAATATGG GTGTGCTGGC
TGGTCAGGCC

1021 TCTCAACTGA ATGCCGTGGT GGACTTGCAA GACAGAAACA CAGAGCTGTC
CTACCAGCTC

1081 TTGCTTGACT CTCTGGGTGA CAGAACCAGG TATTTAGTA TGTGGAATCA
GGCGGTGGAC

Fig. 24A

ITR0048PV

SEQ ID NO: 19

115/153

1141 AGTTATGATC CTGATGTGCG CATTATTGAA AACCATGGTG TGGAGGATGA
ATTGCCAAAC

1201 TATTGCTTCC CCTTGGATGG AGCAGGCACC AATTCGGTTT ACCAAGGTGT
TAAACCAAAA

1261 ACTGACAATG GCAACGATCA GTGGGAAACA GATTCCACAG TTTCAAGTCA
CAATCAGATA

1321 TGCAAAGGCA ATATCTATGC CATGGAGATC AATCTCCAGG CCAACCTGTG
GAGAAGTTTC

1381 CTCTACTCGA ACGTGGCCCT GTACCTGCCC GATTCTTACA AGTACACGCC
GGCCAACATC

1441 ACCCTGCCCC CCAACACCAA CACCTACGAT TACATGAACG GGAGAGTGGT
GCCTCCCTCG

1501 CTGGTGGATG CCTACATCAA CATCGGAGCG CGCTGGTCGC TGGACCCCAT
GGACAACGTC

1561 AATCCCTTCA ACCACCACCG CAATGCGGGG CTGCGCTACC GCTCCATGCT
CCTGGGCAAC

1621 GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAAT TTTTCGCCAT
CAAGAGCCTT

1681 CTGCTCCTGC CCGGGTCCTA CACCTACGAG TGGAATTCC GCAAGGACGT
CAACATGATC

1741 CTGCAGAGCT CCCTCGGCAA CGACCTGCGC ACGGACGGGG CCTCCATCTC
CTTCAACAGC

1801 ATCAACCTCT ACGCCACCTT CTTCCCCATG GCGACAACA CGGCCTCCAC
GCTCGAGGCC

1861 ATGCTGCGCA ACGACACCAA CGACCAGTCC TTCAACGACT ACCTCTCGGC
GGCCAACATG

1921 CTCTACCCCA TCCCGGCCAA CGCCACCAAC GTGCCCATCT CCATCCCCTC
GCGCAACTGG

1981 GCCGCCTTCC GCGGCTGGTC CTTACGCGC CTCAAGACCA AGGAGACGCC
CTCGCTGGGC

2041 TCCGGGTTCG ACCCATACTT CGTCTACTCG GGCTCCATCC CCTACCTCGA
CGGCACCTTC

2101 TACCTCAACC ACACCTTCAA GAAGGTCTCC ATCACCTTCG ATTCCTCCGT
CAGCTGGCCC

2161 GGCAACGACC GGCTCCTGAC GCCCAACGAG TTCGAAATCA AGCGCACCGT
CGACGGCGAG

2221 GGATACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA
GATGCTGGCC

Fig. 24B

ITR0048PV

SEQ ID NO: 19

116/153

2281 CACTACAACA TCGGCTACCA GGGCTTCTAC GTGCCCCGAGG GCTACAAGGA
CCGCATGTAC

2341 TCCTTCTTCC GCAACTTCCA GCCCATGAGC CGCCAGGTGG TGGACGAGGT
CAACTACAAG

2401 GACTACCAGG CCGTCACCTT GGCCTACCAG CACAACAACT CGGGCTTCGT
CGGCTACCTC

2461 GCGCCACCA TGCGCCAGGG CCAGCCCTAC CCCGCCAACT ACCCGTACCC
GCTCATCGGC

2521 AAGAGCGCCG TCACCAGCGT CACCCAGAAA AAGTTCCTCT GCGACAGGGT
CATGTGGCGC

2581 ATCCCCTTCT CCAGCAACTT CATGTCCATG GGCGCGCTCA CCGACCTCGG
GCAGAACATG

2641 CTCTATGCCA ACTCCGCCCA CGCGCTAGAC ATGAATTTTCG AAGTCGACCC
CATGGATGAG

2701 TCCACCCTTC TCTATGTTGT CTTCAAGTC TTCGACGTCG TCCGAGTGCA
CCAGCCCCAC

2761 CGCGGCGTCA TCGAGGCCGT CTACCTGCGC ACCCCCTTCT CGGCCGGTAA
CGCCACCACC

2821 TAA

Fig. 24C

ITR0048PV

SEQ ID NO: 20

117/153

1 ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
61 TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC
CTACTTCAGT
121 CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCACCCACGC ACGATGTGAC
CACCGACCGC
181 AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC
CTACTCGTAC
241 AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACC GCGTGC TGGACATGGC
CAGCACCTAC
301 TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCC TACTC
CGGCACCGCT
361 TACAACAGCC TGGCTCCCAA GGGAGCGCCC AACACTTGCC AGTGGACATA
TACTGATAAC
421 CAAACTGAGA AAACAGCCAC ATATGGAAT GCACCCGTAG AGGGCATTA
CATTACAAAA
481 GATGGCATT C AACTTGAAC TGACAGCGAT GGT CAGGCAA TCTATGCAGA
CGAAACTTAT
541 CAGCCCGAAC CTCAGGTGGG AGATCCTGAA TGGCATGATA CCACAGGTAC
AGAAGAAAAA
601 TATGGAGGCA GAGCGCTTAA ACCTGCCACC GACATGAAAC CTTGCTATGG
CTCTTTTGCC
661 AAGCCAACTA ATGTTAAGGG AGGTCAGGCC AAAAGCAGAA CAAAACTGA
TGGAACTA
721 GAGCCTGATA TTGACATGGC CTTTTTTGAT GGCAGAAATG CAACAACAGC
TGGTTTGACT
781 CCAGAAATTG TTTTGTATAC TGAAAATGTG GATCTGGAAA CTCCAGATAC
CCATATTGTA
841 TACAAGGCAG GCACAGATGA CAGCAGCTCT TCTATCAATT TGGGTCAGCA
GTCCATGCCC
901 AACAGACCCA ACTACATTGG CTT CAGAGAC AAC TTTATCG GGCTCATGTA
CTACAACAGC
961 ACTGGCAATA TGGGTGTACT GGCTGGACAG GCCTCCCAGC TGAATGCTGT
GGTGGACTTG
1021 CAGGACAGAA AACTGAACT GTCCTACCAG CTCTTGCTTG ACTCTCTGGG
TGACAGAACC
1081 AGGTATTTCA GTATGTGGA TCAGGCGGTG GACAGTTATG ACCCCGATGT
GCGCATTTAT

Fig. 25A

ITR0048PV

SEQ ID NO: 20

118/153

1141 GAAATCACG GTGTGGAGGA TGAACCCCC AACTATTGCT TCCCCCTGAA
TGCTGTGGGT

1201 AGAACAAATA GTTATCAGGG AATTAAACCC AATGGAGGCG ATCCAGCTAC
ATGGGCCAAA

1261 GATGAAAGCG TCAATGATTC TAATGAATTG GGCAAGGGCA ATCCTTTTCGC
CATGGAGATC

1321 AACATCCAGG CCAACCTGTG GCGGAACCTC CTCTACGCGA ACGTGGCGCT
GTACCTGCCC

1381 GACTCCTACA AGTACACGCC GGCCAACATC ACGCTGCCCCG CCAACACCAA
CACCTACGAT

1441 TACATGAACG GCCGCGTGGT GGCGCCCTCG CTGGTGGACG CCTACATCAA
CATCGGGGCG

1501 CGCTGGTCGC TGGACCCCAT GGACAACGTC AACCCTTCA ACCACCACCG
CAACGCGGGC

1561 CTGCGCTACC GCTCCATGCT CCTGGGCAAC GGGCGCTACG TGCCCTTCCA
CATCCAGGTG

1621 CCCCAAAAGT TTTTCGCCAT CAAGAGCCTC CTGCTCCTGC CCGGGTCCTA
CACCTACGAG

1681 TGGAACCTCC GCAAGGACGT CAACATGATC CTGCAGAGCT CCCTCGGCAA
CGACCTGCGC

1741 ACGGACGGGG CCTCCATCGC CTTCACCAGC ATCAACCTCT ACGCCACCTT
CTTCCCCATG

1801 GCGCACAACA CCGCCTCCAC GCTCGAGGCC ATGCTGCGCA ACGACACCAA
CGACCAGTCC

1861 TTCAACGACT ACCTCTCGGC GGCCAACATG CTCTACCCCA TCCCGGCCAA
CGCCACCAAC

1921 GTGCCCATCT CCATCCCCTC GCGCAACTGG GCCGCCTTCC GCGGATGGTC
CTTCACGCGC

1981 CTCAAGACCC GCGAGACGCC CTCGCTAGGC TCCGGGTTTCG ACCCCTACTT
CGTCTACTCG

2041 GGCTCCATCC CCTACCTCGA CGGCACCTTC TACCTCAACC ACACCTTCAA
GAAGGTCTCC

2101 ATCACCTTCG ACTCCTCCGT CAGCTGGCCC GGCAACGACC GCCTCCTGAC
GCCAACGAG

2161 TTCGAAATCA AGCGCACCGT CGACGGAGAG GGATACAACG TGGCCCAGTG
CAACATGACC

2221 AAGGACTGGT TCCTGGTCCA GATGCTGGCC CACTACAACA TCGGCTACCA
GGGCTTCTAC

Fig. 25B

ITR0048PV

SEQ ID NO: 20

119/153

2281 GTGCCCAGAGG GCTACAAGGA CCGCATGTAC TCCTTCTTCC GCAACTTCCA
GCCCATGAGC

2341 CGCCAGGTCG TGGACGAGGT CAACTACAAG GACTACCAGG CCGTCACCCT
GGCCTACCAG

2401 CACAACAACCT CGGGCTTCGT CGGCTACCTC GCGCCCACCA TGCGCCAGGG
CCAGCCCTAC

2461 CCCGCCAACT ACCCCTACCC GTCATCGGC AAGAGCGCCG TCGCCAGCGT
CACCCAGAAA

2521 AAGTTCCTCT GCGACCGGGT CATGTGGCGC ATCCCCTTCT CCAGCAACTT
CATGTCCATG

2581 GGC GCGCTCA CCGACCTCGG CCAGAACATG CTCTACGCCA ACTCCGCCCCA
CGCGCTAGAC

2641 ATGAATTTTCG AAGTCGACCC CATGGATGAG TCCACCCTTC TCTATGTTGT
CTTCGAAGTC

2701 TTCGACGTCG TCCGAGTGCA CCAGCCCCAC CGCGGCGTCA TCGAGGCCGT
CTACCTGCGC

2761 ACGCCCTTCT CGGCCGGCAA CGCCACCACC TAA

Fig. 25C

ITR0048PV

SEQ ID NO: 21

120/153

1 ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG
ACAGGACGCT

61 TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCC GCG CCACAGACAC
CTACTTCAGT

121 CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCACGC ACGATGTGAC
CACCGACCGC

181 AGCCAGCGGC TGACGCTGCG CTTCTGCCCC GTGGACCGCG AGGACAACAC
CTACTCGTAC

241 AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACC GCGTGC TGGACATGGC
CAGCACCTAC

301 TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCTACTC
CGGCACCGCC

361 TACAACAGCC TGGCTCCCAA GGGAGCGCCC AACACTTGCC AGTGGACATA
TACTGATAAC

421 CAAACTGAGA AAACAGCCAC ATATGGAAAT GCGCCTGTGC AAGGCATTAG
TATTACAAAA

481 GATGGTATTC AACTTGGAAC TGACACTGAT GATCAGCCCA TTTATGCAGA
TAAAAC TTAT

541 CAACCAGAGC CTCAAGTGGG TGATGCTGAA TGGCATGACA TCACTGGTAC
TGATGAAAAA

601 TATGGAGGCA GAGCTCTCAA GCCTGACACC AAAATGAAGC CCTGCTATGG
TTCTTTTGCC

661 AAGCCTACCA ATAAAGAAGG AGGTCAGGCA AATGTGAAAA CCGAAACAGG
CGGTACCAAA

721 GAATATGACA TTGACATGGC ATTCTTCGAT AATCGAAGTG CAGCTGCGGC
TGGCCTGGCC

781 CCAGAAATTG TTTTGTATAC TGAGAATGTG GATCTGGAAA CTCCAGATAC
TCATATTGTA

841 TACAAGGCAG GCACAGATGA CAGCAGCTCT TCTATCAATT TGGGTCAGCA
GTCCATGCCC

901 AACAGACCCA ACTACATTGG CTTAGAGAC AACTTTATCG GTCTCATGTA
CTACAACAGC

961 ACTGGCAATA TGGGTGTACT GGCTGGTCAG GCCTCCCAGC TGAATGCTGT
GGTGGACTTG

1021 CAGGACAGAA AACTGAACT GTCCTACCAG CTCTTGCTTG ACTCTCTGGG
TGACAGAACC

1081 AGGTATTTTA GSTATGTGGAA TCAGGCGGTG GACAGTTATG ACCCCGATGT
GCGCATTATT

Fig. 26A

ITR0048PV

SEQ ID NO: 21

121/153

1141 GAAAATCACG GTGTGGAGGA TGAAGTCCCT AATTATTGCT TCCCCCTTAA
TGCTGTGGGT

1201 AGAACTGATA CTTACCAGGG AATTAAGGCC AATGGTGCTG ATCAAACCAC
ATGGACCAAA

1261 GATGATACTG TTAATGATGC TAATGAATTG GGCAAGGGCA ATCCTTTTCGC
CATGGAGATC

1321 AACATCCAGG CCAACCTGTG GCGGAAGTTC CTCTACGCGA ACGTGCCCTT
GTACCTGCCC

1381 GACTCCTACA AGTACACGCC GGCCAACATC ACGCTGCCCC CCAACACCAA
CACCTACGAT

1441 TACATGAACG GCCGCGTGGT GCGGCCCTCG CTGGTGGACG CCTACATCAA
CATCGGGGCG

1501 CGCTGGTTCG TGGACCCCAT GGACAACGTC AAGCCCTTCA ACCACCACCG
CAACGCGGGC

1561 CTGCGCTACC GCTCCATGCT CCTGGGCAAC GGGCGCTACG TGCCCTTCCA
CATCCAGGTG

1621 CCCCCAAAGT TCTTCGCCAT CAAGAGCCTC CTGCTCCTGC CCGGGTCTTA
CACCTACGAG

1681 TGGAAGTTCC GCAAGGACGT CAACATGATC CTGCAGAGCT CCCTCGGCAA
CGACCTGCGC

1741 ACGGACGGGG CCTCCATCGC CTTACCAGC ATCAACCTCT ACGCCACCTT
CTTCCCCATG

1801 GCGCACAACA CCGCTCCAC GCTCGAGGCC ATGCTGCGCA ACGACACCAA
CGACCACTCC

1861 TTCAACGACT ACCTCTCGGC GGCCAACATG CTCTACCCCA TCCCGGCCAA
TGCCACCAAC

1921 GTGCCCATCT CCATCCCCTC GCGCAACTGG GCCGCCTTCC GCGGATGGTC
CTTCACGCGC

1981 CTCAAGACCC GCGAGACGCC CTCGCTAGGC TCCGGGTTTC ACCCCTACTT
CGTCTACTCG

2041 GGCTCCATCC CCTACCTCGA CGGCACCTTC TACCTCAACC ACACCTTCAA
GAAGGTCTCC

2101 ATCACCTTCG ACTCCTCCGT CAGCTGGCCC GGCAACGACC GCCTCCTGAC
GCCCAACGAG

2161 TTCGAAATCA AGCGCACCGT CGACGGAGAG GGTACAACG TGGCCCAGTG
CAACATGACC

2221 AAGGACTGGT TCCTGGTCCA GATGCTGGCC CACTACAACA TCGGCTACCA
GGGCTTCTAC

Fig. 26B

ITR0048PV

SEQ ID NO: 21

122/153

2281 GTGCCCAGG GCTACAAGGA CCGCATGTAC TCCTTCTTCC GCAACTTCCA
GCCCATGAGC

2341 CGCCAGGTCG TGGACGAGGT CAACTACAAG GACTACCAGG CCGTCACCCCT
GGCCTACCAG

2401 CACAACAACCT CGGGCTTCGT CGGCTACCTC GCGCCCACCA TGCGCCAGGG
CCAGCCCTAC

2461 CCCGCCAACT ACCCCTACCC GTCATCGGC AAGAGCGCCG TCGCCAGCGT
CACCCAGAAA

2521 AAGTTCTCT GCGACCGGGT CATGTGGCGC ATCCCCTTCT CCAGCAACTT
CATGTCCATG

2581 GGCGCGCTCA CCGACCTCGG CCAGAACATG CTCTACGCCA ACTCCGCCCA
CGCGCTAGAC

2641 ATGAATTTTCG AAGTCGACCC CATGGATGAG TCCACCCTTC TCTATGTTGT
CTTCGAAGTC

2701 TTCGACGTCG TCCGAGTGCA CCAGCCCCAC CGCGGCGTCA TCGAGGCCGT
CTACCTGCGC

2761 ACGCCCTTCT CGGCCGGCAA CGCCACCACC TAA

Fig. 26C

ITR0048PV

SEQ ID NO: 22

123/153

1 ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG
CCAGGACGCC

61 TCGGAGTACC TGAGTCCCGG GCTGGTGCAG TTCGCTCGCG CCACCGAGAG
CTACTTCAGT

121 CTGAGTAACA AGTTTAGGAA CCCCACGGTG GCGCCACGC ACGATGTGAC
CACCAGCCGG

181 TCCCAGCGCC TGACGCTGCG GTTCATCCCC GTGGACCGCG AGGACACCGC
GTACTCGTAC

241 AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACC GCGTGC TGGACATGGC
CTCCACCTAC

301 TTTGACATCC GCGGCGTGCT GGACCGCGGC CCCACCTTCA AGCCCTACTC
CGGCACCGCY

361 TACAACTCCC TGGCCCCCAA GGGCGCTCCC AACTCCTGCG AGTGGGAGCA
AGAGGAAACT

421 CAGGCAGTTG AAGAAGCAGC AGAAGAGGAG GAAGAAGATG CTGACGGTCA
AGCTGAGGAA

481 GAGCAAGCAG CTACCAAAAA GACTCATGTA TATGCTCAGG CTCCCCTTTC
CGGCGAAAAA

541 ATTAGCAAAG ACGGTCTGCA GATAGGAACG GACGCTACAG CAACCGAACA
AAAACCTATT

601 TATGCAGACC CTACATTCCA GCCCGAACCC CAAATCGGGG AGTCCCAGTG
GAATGAGGCA

661 GATGCTACAG TCGCTGGTGG TAGAGTGCTC AAGAAAACCA CTCCCATGAA
ACCATGCTAT

721 GGTTCCTATG CAAGACCCAC GAATGCTAAT GGAGGTCAGG GTGTACTAGC
GGCAAATGCC

781 CAAGGACAGC TAGAATCTCA GGTGAAATG CAATTCTTTT CAACTTCTGA
AAACGCCCCGT

841 AACGAGGCTA ACAACATTCA GCCCAAATTG GTGCTGTATA GCGAGGATGT
GCACATGGAG

901 ACCCCGGATA CACACCTCTC TTACAAGCCC AAAAAAGCG ATGACAATTC
TAAAGTTATG

961 CTGGGCCAAC AGGCCATGCC CAACAGGCCT AATTACATTG GCTTCAGAGA
CAACTTTATC

1021 GGTCTCATGT ACTACAACAG CACTGGCAAC ATGGGAGTGC TTGCAGGTCA
GGCCTCTCAG

1081 TTGAATGCAG TGGTGGACTT GCAAGACAGA AACACAGAAC TGTCTTACCA
GCTCTTGCTT

Fig. 27A

ITR0048PV

SEQ ID NO: 22

124/153

1141 GATTCCATGG GTGACAGAAC CAGATATTTTC TCCATGTGGA ATCAGGCAGT
GGACAGTTAT

1201 GACCCAGATG TCAGAATTAT TGAAAATCAT GGAAGTGAAG ACGAGCTCCC
CAACTATTGT

1261 TTCCCTCTGG GCGGCATAGG GGTAAGTAC ACTTACCAGG CTGTTAAGAC
CAACAATGGC

1321 AATAATGGGG GTCAGGTGAC TTGGACAAAA GATGAACTT TTGCAGAGCG
CAATGAGATA

1381 GGGGTGGGAA ACAATTTTCGC CATGGAGATC AACCTCAATG CCAACCTGTG
GAGGAACTTC

1441 CTGTACTCCA ACGTGGCCCT GTACCTGCCA GACAAGCTTA AGTACAACCC
CTCCAACGTG

1501 GACATCTCTG ACAACCCCAA CACCTACGAT TACATGAACA AGCGAGTGGT
GGCCCCGGGG

1561 CTGGTGGACT GCTACATCAA CCTGGGCGCG CGCTGGTCGC TGGACTACAT
GGACAACGTC

1621 AACCCTTTCA ACCACCACCG CAACGCGGGC CTGCGCTACC GCTCCATGCT
CCTGGGCAAC

1681 GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAGT TCTTTGCCAT
CAAGAACCTC

1741 CTCCTCCTGC CGGGCTCCTA CACCTACGAG TGGAAGTTCA GGAAGGATGT
CAACATGGTC

1801 CTCCAGAGCT CTCTGGGCAA CGATCTCAGG GTGGACGGGG CCAGCATCAA
GTTTCGAGAGC

1861 ATCTGCCTCT ACGCCACCTT CTTCCCCATG GCCCACAACA CCGCCTCCAC
GCTCGAGGCC

1921 ATGCTCAGGA ACGACACCAA CGACCAGTCC TTCAATGACT ACCTCTCCGC
CGCCAACATG

1981 CTCTACCCCA TCCCCGCCAA CGCCACCAAC GTCCCCATCT CCATCCCCTC
GCGCAACTGG

2041 GCGGCCTTCC GCGGCTGGGC CTTACCCGC CTCAAGACCA AGGAGACCCC
CTCCCTGGGC

2101 TCGGGATTCTG ACCCCTACTA CACCTACTCG GGATCCATTC CCTACCTGGA
CGGCACCTTC

2161 TACCTCAACC AACTTTTCAA GAAGGTCTCG GTCACCTTCG ACTCCTCGGT
CAGCTGGCCG

2221 GGCAACGACC GCCTGCTCAC CCCCACGAG TTCGAGATCA AGCGCTCGGT
CGACGGGGAG

Fig. 27B

ITR0048PV

SEQ ID NO: 22

125/153

2281 GGCTACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA
GATGCTGGCC

2341 AACTACAACA TCGGCTACCA GGGCTTCTAC ATCCCAGAGA GCTACAAGGA
CAGGATGTAC

2401 TCCTTCTTCA GGAAGTTCCA GCCCATGAGC CGGCAGGTGG TGGACCAGAC
CAAGTACAAG

2461 GACTACCAGG AGGTGGGCAT CATCCACCAG CACAACAACT CGGGCTTCGT
GGGCTACCTC

2521 GCCCCACCA TGC GCGAGGG ACAGGCCTAC CCCGCCAACT TCCCCTACCC
GCTCATAGGC

2581 AAGACCGCGG TCGACAGCAT CACCCAGAAA AAGTTCCTCT GCGACCGCAC
CCTCTGGCGC

2641 ATCCCCTTCT CCAGCAACTT CATGTCCATG GGTGCGCTCA CGGACCTGGG
CCAGAACCTG

2701 CTCTATGCCA ACTCCGCCCA CGCGCTCGAC ATGACCTTCG AGGTCGACCC
CATGGACGAG

2761 CCCACCCTTC TCTATGTTCT GTTCGAAGTC TTTGACGTGG TCCGGGTCCA
CCAGCCGCAC

2821 CGCGGCGTCA TCGAGACCGT GTACCTGCGC ACGCCCTTCT CGGCCGGCAA
CGCCACCACC

2881 TAA

Fig. 27C

ITR0048PV

SEQ ID NO: 23

126/153

1 ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG
ACAGGACGCT

61 TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCCGCG CCACAGACAC
CTACTTCAGT

121 CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCACGC ACGATGTGAC
CACCAGCCGC

181 AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC
CTACTCGTAC

241 AAAGTGCCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC
CAGCACCTAC

301 TTTGACATCC GCGGCGTGCT GGACCGGGC CCTAGCTTCA AACCTACTC
CGGCACCGCC

361 TACAACAGCC TGGCCCCCAA GGGAGCTCCC AATTCCAGTC AGTGGGAGCA
GACGGAGAAC

421 GGGGGCGGAC AGGCTACGAC TAAAACACAC ACCTATGGAG TTGCCCCAAT
GGGTGGAAT

481 AATATTACAG TCGACGGACT ACAAATTGGA ACTGACGCTA CAGCTGATAC
GGAAAAACCA

541 ATTTATGCTG ATAAAACATT CCAACCTGAG CCTCAGATAG GAGAGGAAAA
CTGGCAAGAA

601 ACTGAAAGCT TTTATGGCGG TAGGGCTCTT AAGAAAGACA CAAACATGAA
GCCTTGTTAT

661 GGCTCATTTG CCAGACCTAC CAATGAAAAG GGAGGTCAAG CTAAACTTAA
AGTTGGAGCT

721 GATGGGCTGC CGACCAAAGA ATTTGACATA GACCTAGCAT TCTTTGATAC
TCCTGGTGGC

781 ACTGTGACCG GAGGTACAGA GGAGTATAAA GCAGATATTG TTATGTATAC
CGAAAACACG

841 TATCTGGAAG CTCCAGACAC ACATGTGGTG TATAAACAG GCAAGGATAA
CACAAGTTCT

901 AAAATTAACC TGGTCCAGCA GTCTATGCCC AACAGGCCCA ACTACATTGG
GTTTAGGGAC

961 AACTTTATTG GGCTCATGTA TTACAACAGC ACTGGCAATA TGGGTGTGCT
GGCCGGTCAG

1021 GCTTCTCAGT TGAATGCTGT GGTTGACTTG CAAGACAGAA AACTGAACT
GTCTTACCAG

1081 CTCTTGCTTG ACTCTTTGGG TGACAGAACC AGGTATTTC AATGTGGAA
TCAGGCGGTG

Fig. 28A

ITR0048PV

SEQ ID NO: 23

127/153

1141 GACAGTTATG ATCCTGATGT GCGCATTATT GAAAACCATG GTGTGGAAGA
TGAACCTCCC

1201 AACTATTGCT TCCCCCTGGA TGGGTCTGGC ACTAACGCCG CTTACCAAGG
TGTGAAAGTA

1261 AAAAATGGTC AAGATGGTGA TGTGAGAGC GAATGGGAAA AAGATGATAC
TGTCGCAGCT

1321 CGAAATCAAT TATGCAAGGG CAACATTTT GCCATGGAGA TCAATCTCCA
GGCCAACCTG

1381 TGGAGAAGTT TTCTCTACTC GAACTGGCC CTGTACCTGC CCGATTCTTA
CAAGTACACG

1441 CCGGCCAACA TCACCCTGCC CACCAACACC AACACCTACG ATTACATGAA
CGGGAGAGTG

1501 GTGCCTCCCT CGCTGGTGA CGCTACATC AACATCGGG CGCGCTGGTC
GCTGGACCCC

1561 ATGGACAACG TCAATCCCTT CAACCACCAT CGCAACGCGG GGCTGCGCTA
CCGCTCCATG

1621 CTCCTGGGCA ACGGGCGCTA CGTGCCCTTC CACATCCAGG TGCCCCAGAA
ATTTTTCGCC

1681 ATTAAGAGCC TCCTGCTCCT GCCCGGTCC TACACCTACG AGTGGAAGTT
CCGCAAGGAC

1741 GTCAACATGA TCCTGCAGAG CTCCTCGGC AACGACCTGC GCACGGACGG
GGCCTCCATC

1801 TCCTTCACCA GCATCAACCT CTACGCCACC TTCTTCCCCA TGGCGCACAA
CACCGCCTCC

1861 ACGCTCGAGG CCATGCTGCG CAACGACACC AACGACCAGT CCTTCAACGA
CTACCTCTCG

1921 GCGGCCAACA TGCTCTACCC CATCCCGGCC AACGCCACCA ACGTGCCCAT
CTCCATCCCC

1981 TCGCGCAACT GGGCCGCCTT CCGCGGCTGG TCCTTCACGC GCCTCAAGAC
CAAGGAGACG

2041 CCCTCGCTGG GCTCCGGGT CGACCCCTAC TTCGTCTACT CGGGCTCCAT
CCCCTACCTC

2101 GACGGCACCT TCTACCTCAA CCACACCTTC AAGAAGGTCT CCATCACCTT
CGACTCCTCC

2161 GTCAGCTGGC CCGCAACGA CCGGCTCCTG ACGCCCAACG AGTTCGAAAT
CAAGCGCACC

2221 GTCGACGGCG AGGGCTACAA CGTGGCCCAG TGCAACATGA CCAAGGACTG
GTTCTGCTG

2281 CAGATGCTGG CCCACTACAA CATCGGCTAC CAGGGCTTCT ACGTGCCCGA
GGGCTACAAG

Fig. 28B

ITR0048PV

SEQ ID NO: 23

128/153

2341 GACCGCATGT ACTCCTTCTT CCGCAACTTC CAGCCCATGA GCCGCCAGGT
CGTGGACGAG

2401 GTCAACTACA AGGACTACCA GGCCGTCACC CTGGCCTACC AGCACAACAA
CTCGGGCTTC

2461 GTCGGCTACC TCGCGCCAC CATGCGCCAG GGCCAGCCCT ACCCCGCCAA
CTACCCCTAC

2521 CCGCTCATCG GCAAGAGCGC CGTCGCCAGC GTCACCCAGA AAAAGTTCCT
CTGCGACCGG

2581 GTCATGTGGC GCATCCCCCTT CTCCAGCAAC TTCATGTCCA TGGGCGCGCT
CACCGACCTC

2641 GGCCAGAACA TGCTCTACGC CAACTCCGCC CACGCGCTAG ACATGAATTT
CGAAGTCGAC

2701 CCCATGGATG AGTCCACCCT TCTCTATGTT GTCTTCGAAG TCTTCGACGT
CGTCCGAGTG

2761 CACCAGCCCC ACCGCGGCGT CATCGAGGCC GTCTACCTGC GCACCCCTT
CTCGGCCGGT

2821 AACGCCACCA CCTAA

Fig. 28C

ITR0048PV

SEQ ID NO: 24

129/153

1 ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG
CCAGGACGCC

61 TCNGAGTACC TGAGCCCCGG GCTGGTGCAG TTCGCCCGCG CCACCGAGAG
CTACTTCAGC

121 CTGAGTAACA AGTTTAGGAA CCCACGGTG GCGCCACGC ACGATGTGAC
CACCGACCGG

181 TCTCAGCGCC TGACGCTGCG GTTCATTCCC GTGGACCGCG AGGACACCGC
GTACTCGTAC

241 AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC
CTCCACCTAC

301 TTTGACATCC GCGGGGTGCT GGACCGGGT CCCACTTTCA AGCCCTACTC
TGGCACCGCC

361 TACAACTCCC TGGCCCCCAA GGGCGTCCC AACTCCTGCG AGTGGGAGCA
AGAGGAAACT

421 CAGGCAGTTG AAGAAGCAGC AGAAGAGGAA GAAGAAGATG CTGACGGTCA
AGCTGAGGAA

481 GAGCAAGCAG CTACCAAAAA GACTCATGTA TATGCTCAGG CTCCCCTTTC
TGGCGAAAAA

541 ATTAGTAAAG ATGGTCTGCA AATAGGAACG GACGCTACAG CTACAGAACA
AAAACCTATT

601 TATGCAGACC CTACATTCCA GCCCGAACCC CAAATCGGGG AGTCACAGTG
GAATGAGGCA

661 GATGCTACAG TCGCCGGCGG TAGAGTGCTA AAGAAATCTA CTCCCATGAA
ACCATGCTAT

721 GGTTCCTATG CAAGACCCAC AAATGCTAAT GGAGGTCAGG GTGTACTAAC
GGCAAATGCC

781 CAGGGACAGC TAGAATCTCA GGTTGAAATG CAATTCTTTT CAACTTCTGA
AAACGCCCCG

841 AACGAGACTA ACAACATTCA GCCCAAATTG GTGCTGTATA GTGAGGATGT
GCACATGGAG

901 ACCCCGGATA CGCACCTTTC TTACAAGCCC GCAAAAAGCG ATGACAATTC
AAAAATCATG

961 CTGGGTCAGC AGTCCATGCC CAACAGACCT AATTACATCG GCTTCAGAGA
TAACTTTATC

1021 GGCCTCATGT ATTACAATAG CACTGGCAAC ATGGGAGTGC TTGCAGGTCA
GGCCTCTCAG

1081 TTGAATGCAG TGGTGGACTT GCAAGACAGA AACACAGAAC TGTCTTACCA
GCTCTTGCTT

Fig. 29A

ITR0048PV

SEQ ID NO: 24

130/153

1141 GATTCCATGG GTGACAGAAC CAGATACTTT TCCATGTGGA ATCAGGCAGT
GGACAGTTAT

1201 GACCCAGATG TTAGAATTAT TGAAAATCAT GGAAGTGAAG ACGAGCTCCC
CAACTATTGT

1261 TTCCCTCTGG GTGGCATAGG GGTAAGTAC ACTTACCAGG CTGTTAAAC
CAACAATGGC

1321 AATAACGGGG GCCAGGTGAC TTGGACAAAA GATGAACTT TTGCAGATCG
CAATGAAATA

1381 GGGGTGGGAA ACAATTTGCG TATGGAGATA AACCTCAGTG CCAACCTGTG
GAGAACTTC

1441 CTGTACTCCA ACGTGGCGCT GTACCTACCA GACAAGCTTA AGTACAACCC
CTCCAATGTG

1501 GACATCTCTG ACAACCCCAA CACCTACGAT TACATGAACA AGCGAGTGGT
GGCCCCGGGG

1561 CTGGTGGACT GCTACATCAA CCTGGGCGCG CGCTGGTCGC TGGACTACAT
GGACAACGTC

1621 AACCCTTCA ACCACCACCG CAATGCGGGC CTGCGCTACC GCTCCATGCT
CCTGGGCAAC

1681 GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAGT TCTTTGCCAT
CAAGAACCTC

1741 CTCCTCTGCG CGGGCTCCTA CACCTACGAG TGGAAGTTCA GGAAGGATGT
CAACATGGTC

1801 CTCCAGAGCT CTCTGGGTAA CGATCTCAGG GTGGACGGGG CCAGCATCAA
GTTGAGAGC

1861 ATCTGCCTCT ACGCCACCTT CTTCCTCATG GCCCACAACA CGGCCTCCAC
GCTCGAGGCC

1921 ATGCTCAGGA ACGACACCAA CGACCAGTCC TTCAATGACT ACCTCTCCGC
CGCCAACATG

1981 CTCTACCCCA TACCCGCCAA CGCCACCAAC GTCCCCATCT CCATCCCCTC
GCGCAACTGG

2041 GCGGCCTTCC GCGGCTGGGC CTTACCCGC CTCAAGACCA AGGAGACCCC
CTCCCTGGGC

2101 TCGGGATTCTG ACCCTACTA CACCTACTCG GGCTCCATTC CCTACCTGGA
CGGCACCTTC

2161 TACCTCAACC AACTTTCAA GAAGGTCTCG GTCACCTTCG ACTCCTCGGT
CAGCTGGCCG

2221 GGCAACGACC GTCTGCTCAC CCCCACGAG TTCGAGATCA AGCGCTCGGT
CGACGGGGAG

Fig. 29B

ITR0048PV

SEQ ID NO: 24

131/153

2281 GGCTACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA
GATGCTGGCC

2341 AACTACAACA TCGGCTACCA GGGCTTCTAC ATCCCAGAGA GCTACAAGGA
CAGGATGTAC

2401 TCCTTCTTCA GGAAC TTCCA GCCCATGAGC CGGCAGGTGG TGGACCAGAC
CAAGTACAAG

2461 GACTACCAGG AGGTGGGCAT CATCCACCAG CACAACAAC TCGGGCTTCGT
GGGCTACCTC

2521 GCCCCACCA TGC GCGAGGG ACAGGCCTAC CCCGCCAACT TCCCCTATCC
GCTCATAGGC

2581 AAGACCGCGG TCGACAGCAT CACCCAGAAA AAGTTCCTCT GCGACCGCAC
CCTCTGGCGC

2641 ATCCCCTTCT CCAGCAACTT CATGTCCATG GGTGCGCTCT CGGACCTGGG
CCAGAACTTG

2701 CTCTACGCCA ACTCCGCCCCA CGCCCTCGAC ATGACCTTCG AGGTCGACCC
CATGGACGAG

2761 CCCACCCTTC TCTATGTTCT GTTCGAAGTC TTTGACGTGG TCCGGGTCCA
CCAGCCGCAC

2821 CGCGGCGTCA TCGAGACCGT GTACCTGCGT ACGCCCTTCT CGGCCGGCAA
CGCCACCACC

2881 TAA

Fig. 29C

ITR0048PV

SEQ ID NO: 25

132/153

1 ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG
CCAGGACGCC

61 TCGGAGTACC TGAGCCCCGG GCTGGTGCAG TTCGCCCCGG CCACCGAGAG
CTACTTCAGT

121 CTGAGTAACA AGTTTAGGAA CCCCACGGTG GCGCCACGC ACGATGTGAC
CACCGACCGG

181 TCCCAGCGCC TGACGCTGCG GTTCATCCCC GTGGACCGCG AGGACACCGC
GTACTCGTAC

241 AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC
CTCCACCTAC

301 TTTGACATCC GCGGCGTGCT GGACCGCGGC CCCACCTTCA AGCCCTACTC
CGGCACCGCC

361 TACAAC TCCC TGGCCCCCAA GGGCGCTCCC AACTCTTG TG AGTGGGAGCA
ATTAGAAGAA

421 GCCCAGGCCG CTTTGGAAGA CGAAGAATTA GAAGATGAAG ACGAGGAACC
ACAGGATGAG

481 GCGCCTGTGA AAAAGACCCA TGTATACGCT CAGGCTCCCC TTTCTGGAGA
AGAAATTACT

541 AAAGACGGTT TGCAAATAGG GTCAGATAAC ACAGAAGCTC AGTCTAAGCC
TATATATGCA

601 GACCCTACAT TCCAGCCCGA ACCCCAAATC GGGGAGTCCC AGTGGAACGA
GGCAGATGCT

661 ACAGTCGCTG GTGGTAGAGT GCTCAAGAAA ACCACTCCCA TGAAACCATG
CTATGGTTCC

721 TATGCAAGAC CCACGAATGC TAATGGAGGT CAGGGTGTGC TGGTGGCTGA
TGATAAGGGG

781 GTCCTTCAAT CTAAAGTTGA ATTGCAATTT TTTTCAAATA CTACTACTCT
TAATCAGCGG

841 GAGGGTAATG ATACAAAACC AAAAGTAGTG CTGTATAGCG AGGATGTGCA
CATGGAAACA

901 CCAGACACCC ACATTTCTTA CAAGCCCACA AAAAGCGATG ACAATTCTAA
AGTTATGCTG

961 GGCCAACAGT CCATGCCCAA CAGGCCTAAT TACATCGGCT TCAGAGACAA
CTTTATCGGT

1021 CTCATGTACT ACAACAGCAC TGGCAACATG GGAGTGCTTG CAGGTCAGGC
CTCTCAGTTG

1081 AATGCAGTGG TGGACTTGCA AGACAGAAAC ACAGAACTGT CCTACCAGCT
CTTGCTTGAT

Fig. 30A

ITR0048PV

SEQ ID NO: 25

133/153

1141 TCCATGGGTG ACAGAACCAG ATATTTCTCC ATGTGGAATC AGGCAGTGGA
CAGTTATGAC

1201 CCGGATGTCA GAATTATTGA AAATCATGGA ACCGAAGACG AGCTCCCCAA
CTATTGTTTT

1261 CCTCTGGGTG GCATAGGGGT AACTGACACT TACCAGGTCA TTAAACTAA
TGGCAATGGT

1321 CAAGCAGACC CAACCTGGGA AAAAGATACA GAGTTTGCAG ACCGCAATGA
AATAGGGGTG

1381 GGAAACAATT TCGCCATGGA GATCAACCTC AATGCCAACC TGTGGAGGAA
CTTCCTGTAC

1441 TCCAACGTGG CCCTGTACCT GCCAGACAAG CTTAAGTACA ACCCCTCCAA
CGTGGACATC

1501 TCTGACAACC CCAACACCTA CGATTACATG AACAAAGCGAG TGGTGGCCCC
GGGGCTGGTG

1561 GACTGCTACA TCAACCTGGG CGCGCGCTGG TCGCTGGACT ACATGGACAA
CGTCAACCCC

1621 TTCAACCACC ACCGCAACGC GGGCCTGCGC TACCGCTCCA TGCTCCTGGG
CAACGGGCGC

1681 TACGTGCCCT TCCACATCCA GGTGCCCCAG AAGTTCTTTG CCATCAAGAA
CCTCCTCCTC

1741 CTGCCGGGCT CCTACACCTA CGAGTGGAAC TTCAGGAAGG ATGTCAACAT
GGTCCTCCAG

1801 AGCTCTTTGG GCAACGATCT CAGGGTGGAC GGGGCCAGCA TCAAGTTCGA
GAGCATCTGC

1861 CTCTACGCCA CCTTCTTCCC CATGGCCAC AACACCGCCT CCACGCTCGA
GGCCATGCTC

1921 AGGAACGACA CCAACGACCA GTCCTTCAAT GACTACCTCT CCGCCGCCAA
CATGCTCTAC

1981 CCCATCCCCG CCAACGCCAC CAACGTCCCT ATCTCCATCC CCTCGCGCAA
CTGGGCGGCC

2041 TTCCGCGGCT GGGCCTTCAC CCGCCTCAAG ACCAAGGAGA CACCCTCCCT
GGGCTCGGGA

2101 TTCGACCCCT ACTACACCTA CTCGGGATCC ATTCCCTACC TGGACGGCAC
CTTCTACCTC

2161 AACCACACTT TCAAGAAGGT CTCGGTCACC TTCGACTCCT CGGTCAGCTG
GCCGGGCAAC

2221 GACCGCCTGC TCACCCCCAA CGAGTTCGAG ATCAAGCGCT CGGTGACGG
GGAGGGCTAC

Fig. 30B

ITR0048PV

SEQ ID NO: 25

134/153

2281 AACGTGGCCC AGTGCAACAT GACCAAGGAC TGGTTCCTGG TCCAGATGCT
GGCCAACTAC

2341 AACATCGGCT ACCAGGGCTT CTACATCCCA GAGAGCTACA AGGACAGGAT
GTACTCCTTC

2401 TTCAGGAACT TCCAGCCCAT GAGCCGGCAG GTGGTGGACC AAACCAAGTA
CAAGGACTAC

2461 CAGGAGGTGG GCATCATCCA CCAGCACAAC AACTCGGGCT TCGTGGGCTA
CCTCGCCCCC

2521 ACCATGCGCG AGGGACAGGC CTACCCCGCC AACTTCCCCT ACCCGCTCAT
AGGCAAGACC

2581 GCGGTCGACA GCATCACCCA GAAAAAGTTC CTCTGCGACC GCACCCTCTG
GCGCATCCCC

2641 TTCTCCAGCA ACTTCATGTC CATGGGTGCG CTCACGGACC TGGGCCAGAA
CCTGCTCTAT

2701 GCCAACTCCG CCCACGCGCT CGACATGACC TTCGAGGTCG ACCCATGGA
CGAGCCCACC

2761 CTTCTCTATG TTCTGTTCGA AGTCTTTGAC GTGGTCCGGG TCCACCAGCC
GCACCGCGGC

2821 GTCATCGAGA CCGTGTACCT GCGCACGCCC TTCTCGGCCG GCAACGCCAC
CACCTAA

Fig. 30C

1					50
had12	MATPSMMPQW	SYMHIAGQDA	SEYLSPLVQ	FARATDTYFT	LGNKFRNPTV
had3	MATPSMMPQW	AYMHIAGQDA	SGYLSPLVQ	FARATDTYFS	MGNKFRNPTV
had7	MATPSMMPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	MGNKFRNPTV
had11	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFN	LGNKFRNPTV
had21	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFN	LGNKFRNPTV
had34	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFN	LGNKFRNPTV
had35	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFN	LGNKFRNPTV
C1	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFN	LGNKFRNPTV
had1	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATETYFS	LNNKFRNPTV
had2	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATETYFS	LNNKFRNPTV
had5	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATETYFS	LNNKFRNPTV
Chad3	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATESYFS	LSNKFRNPTV
Chad11	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATESYFS	LSNKFRNPTV
Chad17	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATESYFS	LSNKFRNPTV
Chad19	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATESYFS	LSNKFRNPTV
Chad20	MATPSMMPQW	SYMHISGQDA	SEYLSPLVQ	FARATESYFS	LSNKFRNPTV
had48	MATPSMMPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Chad4	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Chad5	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Chad7	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Chad16	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Pan6	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
had4	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
had16	MATPSMMPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	MGNKFRNPTV
Chad6	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Chad9	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Chad10	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
C6	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Pan5	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
Pan7	MATPSMLPQW	AYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
had41	MATPSMMPQW	SYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
had40	MATPSMMPQW	SYMHIAGQDA	SEYLSPLVQ	FARATDTYFS	LGNKFRNPTV
51					100
had12	APTHDVTDDR	SQRLTLRFVP	VDREDTYSY	KARFTLAVGD	NRVLDMASSY
had3	APTHDVTDDR	SQRLMLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTF
had7	APTHDVTDDR	SQRLMLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTF
had11	APTHDVTDDR	SQRLMLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTF
had21	APTHDVTDDR	SQRLMLRFVP	VDREDNTYAY	KVRYTLAVGD	NRVLDMASTF
had34	APTHDVTDDR	SQRLMLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTF
had35	APTHDVTDDR	SQRLMLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTF
C1	APTHDVTDDR	SQRLMLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTF
had1	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
had2	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
had5	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
Chad3	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
Chad11	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
Chad17	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
Chad19	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
Chad20	APTHDVTDDR	SQRLTLRFIP	VDREDTAYSY	KARFTLAVGD	NRVLDMASTY
had48	APTHDVTDDR	SQRLTLRFVP	VDREDTYSY	KARFTLAVGD	NRVLDMASTY
Chad4	APTHDVTDDR	SQRLTLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTY
Chad5	APTHDVTDDR	SQRLTLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTY
Chad7	APTHDVTDDR	SQRLTLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTY
Chad16	APTHDVTDDR	SQRLTLRFVP	VDREDNTYSY	KVRYTLAVGD	NRVLDMASTY

Fig. 31A

```

Pan6  APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
hAd4  APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
hAd16 APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
ChAd6 APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
ChAd9 APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
ChAd10 APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
C68   APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
Pan5  APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
Pan7  APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
hAd41 APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRTTLAVGD NRVLDMASTY
hAd40 APTHDVTTDR SQRLTLRFVP VDREETAYSY KVRTTLAVGD NRVLDMASTY

```

```

101                                     150
hAd12  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NASQWSD... ..
hAd3   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTSQWIVT TN GDNA.....
hAd7   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTSQWIVT AG EERA.....
hAd11  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTSQWIAEGV KNTTGEEHVT
hAd21  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTSQWIAEGV KKEDGGSDEE
hAd34  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NASQWLDKGV TSTGLVDDGN
hAd35  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NASQWIAKGV PTAAAAGNGE
C1     FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTSQWLDKGV TTTDNNTENG
hAd1   FDIRGVLD RG PTFKPYSGTA YNALAPKGAP NSCEWEQEEP TQEMAELEED
hAd2   FDIRGVLD RG PTFKPYSGTA YNALAPKGAP NSCEWEQTED SGRAVAEDEE
hAd5   FDIRGVLD RG PTFKPYSGTA YNALAPKGAP NPCEWDEAAT ALEINLEED
ChAd3  FDIRGVLD RG PTFKPYSGTA YNSLAPKGAP NSCEWEQ.EE TQAVEEAAEE
ChAd11 FDIRGVLD RG PTFKPYSGTA YNSLAPKGAP NSCEWEQ.EE TQAVEEAAEE
ChAd17 FDIRGVLD RG PTFKPYSGTA YNSLAPKGAP NSCEWEQ.EE TQAVEEAAEE
ChAd19 FDIRGVLD RG PTFKPYSGTA YNSLAPKGAP NSCEWEQLEE AQAALEDEEL
ChAd20 FDIRGVLD RG PTFKPYSGTA YNSLAPKGAP NPCEWDEAAT ALDIDLNAED
hAd48  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NPSQWEEKKN GGGS.....
hAd4   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NSSQWEQKKT GNNA.....
ChAd5  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NSSQWEQKKT GNNA.....
ChAd7  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NSSQWEQKKT GKNA.....
ChAd16 FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NSSQWEQKTEN G.....
Pan6   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NSSQWEQAKT G.....
hAd4   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWKD... ..
hAd16  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWKD... ..
ChAd6  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTSQWITKDN .....
ChAd9  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWITYTDN .....
ChAd10 FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWITYTDN .....
C68    FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWITYKAD GE.....
Pan5   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWITYKAD G.....
Pan7   FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NTCQWITYKAG D.....
hAd41  FDIRGVLD RG PSFKPYSGTA YNSLAPKTAP NPCEWKD... ..
hAd40  FDIRGVLD RG PSFKPYSGTA YNSLAPKGAP NPSQWTN... ..

```

```

151                                     200
hAd12  ..... .NAKLNTFAQ APYLSD...T ITAADGIKVG
hAd3   ..... VTTTTNTFGI ASMKGG...N ITKE.GLQIG
hAd7   ..... VTTTTNTFGI ASMKGD...N ITKE.GLEIG
hAd11  EEE..... TTTTTYTFGN APVKAEA..E ITKE.GLPVG
hAd21  EEK..... .NLTTYTFGN APVKAEG.GD ITKDKGLPIG
hAd34  DDD..... .GEE AKKATYTFGN APVKAEA..E ITKD.GLPVG
hAd35  EEH..... .ETE EKTATYTFAN APVKAEA..Q ITKE.GLPVG
C1     DE..... EDEVAEEGEE EKQATYTFGN APVKAEA..E ITKE.GLPVG
hAd1   EEEAE EEEAE EAEAPQADQ KVKKTHVYAO APLAGE...K ITAN.GLQIV

```

Fig. 31B

```

hAd2   EEE.DEDEEE EEEEEQNARDQ ATKKTHVYAO APLSGE...T ITKS.GLQIG
hAd5   DDN....EDE VDEQAE.... .QQKTHVFGQ APYSGI...N ITKE.GIQIG
ChAd3   EE.....ED ADGQAEEDQA ATKKTHVYAO APLSGE...K ISKD.GLQIG
ChAd11  EE.....ED ADGQAEEDQA ATKKTHVYAO APLSGE...K ISKD.GLQIG
ChAd17  EE.....ED ADGQAEEDQA ATKKTHVYAO APLSGE...K ISKD.GLQIG
ChAd19  ED..... .EDEEPQDEA PVKKTHVYAO APLSGE...E ITKD.GLQIG
ChAd20  DE..... .ESDEAQGEA DQKTHVFGQ APYSGQ...N ITKE.GIQIG
hAd48  ..... .DA NQMOTHTFGV AAMGGI...E ITAK.GLQIG
ChAd4   ..... .N GDTENVTYGV AAMGGI...D IDKN.GLQIG
ChAd5   ..... .N GDTENVTYGV AAMGGI...D IDKN.GLQIG
ChAd7   ..... .N GDTENVTYGV AAMGGI...D IDKN.GLQIG
ChAd16  ..... .GGQ ATTKHTYGV APMGGT...N ITVD.GLQIG
Pan6    ..... .NG GTMETHYGV APMGGE...N ITKD.GLQIG
hAd4    ..... .ANSKMHTFGV AAMPVGTGKK IEAD.GLPIR
hAd16   ..... .SDSKMHTFGV AAMPVGTGKK IEAD.GLPIG
ChAd6   ..... .GTDKTSFGN APVRGL...D ITEE.GLQIG
ChAd9   ..... .QTEKTATYGN APVEGI...N ITKD.GIQLG
ChAd10  ..... .QTEKTATYGN APVQGI...S ITKD.GIQLG
C68     ..... .T ATEKTYTYGN APVQGI...N ITKD.GIQLG
Pan5    ..... .DT GTEKTYTYGN APVQGI...S ITKD.GIQLG
Pan7    ..... .T DTEKTYTYGN APVQGI...S ITKD.GIQLG
hAd41   ..... .NNKIKVRGQ APFIGT...N INKDNQIQIG
hAd40   ..... .QNKTNSFGQ APYIGQ...K ITNQ.GVQVG

```

201

250

```

hAd12   TDTAQA...G AAVYANKTYQ PEPQVGPSEW NTSIE.NVKA GGRALKQTTA
hAd3    KDITTTTEGEE KPIYADKTYQ PEPQVGEESW TDTDGTNEKF GGRALKPATN
hAd7    KDITAD...N KPIYADKTYQ PEPQVGEESW TDTDGTNEKF GGRALKPATK
hAd11   LEVSDE..ES KPIYADKTYQ PEPQLGDETW TDLDGKTEKY GGRALKPDTK
hAd21   SEITDG..EA KPIYADKLYQ PEPQVGDETW TDTDGTTEKY GGRALKPETK
hAd34   LEVSTE..GP KPIYADKLYQ PEPQVGDETW TDLDGKTEEY GGRVLKPETK
hAd35   LEISAE..NES KPIYADKLYQ PEPQVGDETW TDLDGKTEEY GGRALKPDTK
C1      LEVPSEG.DP KPIYADKLYQ PEPQVGEESW TDTDGTDEKY GGRALKPETK
hAd1    SDTQTE...G NPVFADPTYQ PEPQVGESQW NEAEA..TAS GGRVLKKTTP
hAd2    SDNAET...QA KPVIADPSYQ PEPQIGESQW NEADA..NAA GGRVLKKTTP
hAd5    VEGQ..... TPKYADKTFQ PEPQIGESQW YETEI..NHA AGRVLKKTTP
ChAd3   TDATA...EQ KPIYADPTFQ PEPQIGESQW NEADA..TVA GGRVLKKTTP
ChAd11  TDATA...EQ KPIYADPTFQ PEPQIGESQW NEADA..TVA GGRVLKKTTP
ChAd17  TDATA...EQ KPIYADPTFQ PEPQIGESQW NEADA..TVA GGRVLKKTTP
ChAd19  SDNTEA..QS KPIYADPTFQ PEPQIGESQW NEADA..TVA GGRVLKKTTP
ChAd20  IDAASQ..AQ TPVIADKTFQ PEPQVGESQW NETEI..SYG AGRVLKKTTL
hAd48   IDATKEEDNG KEIYADKTFQ PEPQIGEENW QSD...NYY GGRAIKKETK
ChAd4   TDDTKD..DD NEIYADKTYQ PEPQIGEENW QETY...SYY GGRALKKDTK
ChAd5   TDDTKD..DD NEIYADKTYQ PEPQIGEENW QETY...SYY GGRALKKDTK
ChAd7   TDDTKD..GD NEIYADKTYQ PEPQIGEENW QETY...SYY GGRALKKDTK
ChAd16  TDATA...TE KPIYADKTFQ PEPQIGEENW QETE...SFY GGRALKKDTN
Pan6    TDVTAN...QN KPIYADKTFQ PEPQVGEENW QETE...NFY GGRALKKDTN
hAd4    IDSTSG..TD TVIYADKTFQ PEPQVGNDSW VDTNDAEEKY GGRALKDTTN
hAd16   IDSTSG..TD TVIYADKTFQ PEPQVGNASW VDANGTEEKY GGRALKDTTK
ChAd6   PDESGG..ES KKIFADKTYQ PEPQLGDEEW HDTIGAEDKY GGRALKPATN
ChAd9   TDS DG..... QAIYADETYQ PEPQVGDP EW HDITGTDEKY GGRALKPATD
ChAd10  TDTDD..... QPIYADKTYQ PEPQVGDAEW HDITGTDEKY GGRALKPDTK
C68     TDTDD..... QPIYADKTYQ PEPQVGDAEW HDITGTDEKY GGRALKPDTK
Pan5    TDTDD..... QPIYADKTYQ PEPQVGDAEW HDITGTDEKY GGRALKPDTK
Pan7    TDS DG..... QAIYADETYQ PEPQVGDAEW HDITGTDEKY GGRALKPDTK
hAd41   TDTTN..... QPIYADKTYQ PEPQVGQTQW NSEVGAAQKV AGRVLKDTTP
hAd40   SDSNN..... RDVFADKTYQ PEPQVGQTQW NINPM..QNA AGRILKQTTTP

```

Fig. 31C

	251		300
hAd12	MQPCYGSYAR	PTNEHGGQS.KDDNIE	LKFFDSANNA
hAd3	MKPCYGSFAR	PTNIKGGQAK NRKVKPTTEG	GVETEEDID MEFFDGRDAV
hAd7	MKPCYGSFAR	PTNIKGGQAK NRKVKP.TEG	DVETEEDID MEFFDGREAA
hAd11	MKPCYGSFAK	PTNVKGGQAK QKTTEQPN..	..QKVEYDID MEFFDAASQK
hAd21	MKPCYGSFAK	PTNVKGGQAK QKTTEQPQ..	..NQVEYDID MNFFDEASQK
hAd34	MKPCYGSFAK	PTNIKGGQAK VKPKEDDG..	..TNNIEYDID MNFFDLRSQR
hAd35	MKPCYGSFAK	PTNVKGGQAK QKTTEQP...	..NQKVEYDID MEFFDAASQR
C1	MKPCYGSFAK	PTNVKGGQAK VKKVEEG...	..KVEYDID MNFFDLRSQK
hAd1	MKPCYGSYAR	PTNKNGGQGI LVANNQG...	..ALESKVE MQFFAPSGTA
hAd2	MKPCYGSYAR	PTNPFGGQSV LVPDEKG...	..VPLPKVD LQFFSNTTSL
hAd5	MKPCYGSYAK	PTNENGGQGI LVKQONG...	..KLESQVE MQFFSTTEAT
ChAd3	MKPCYGSYAR	PTNANGGQGV LTANAQG...	..QLESQVE MQFFSTSENA
ChAd11	MKPCYGSYAR	PTNANGGQGV LAANAQG...	..QLESQVE MQFFSTSENA
ChAd17	MKPCYGSYAR	PTNANGGQGV LTANAQG...	..QLESQVE MQFFSTSENA
ChAd19	MKPCYGSYAR	PTNANGGQGV LVADDKG...	..VLQSKVE LQFFSNTTTL
ChAd20	MKPCYGSYAR	PTNENGGQGI LLEQDG....	..KKESQVE MQFFSTTQAA
hAd48	MKPCYGSFAR	PTNEKGGQAK FKTPEKEGE..	..EPKELDID LNFFDIPSTG
ChAd4	MKPCYGSFAR	PTNVKGGQAK IKTDGD....	..VKSFDID LAFFDIPNSG
ChAd5	MKPCYGSFAR	PTNVKGGQAK IKTDGD....	..VKSFDID LAFFDIPNSG
ChAd7	MKPCYGSFAR	PTNVKGGQAK IKTDGD....	..VKSFDID LAFFDIPNSG
ChAd16	MKPCYGSFAR	PTNEKGGQAK LKVGADG...	..LPTKEFDID LAFFDTPGGT
Pan6	MKPCYGSYAR	PTNEKGGQAK LKVGDDG...	..VPTKEFDID LAFFDTPGGT
hAd4	MNPCYGSFAK	PTNKEGGQAN LKDSETA...	..ATTPNYDID LAFFDGKNIV
hAd16	MKPCYGSFAK	PTNKEGGQAN LKDSETA...	..ATTPNYDID LAFFDNKNIA
ChAd6	MKPCYGSFAK	PTNAKGGQAK SRTKDDG...	..TTEPDID MAFFDDRSQQ
ChAd9	MKPCYGSFAK	PTNVKGGQAK SRTKTDG...	..TTEPDID MAFFDGRNAT
ChAd10	MKPCYGSFAK	PTNKEGGQAN VKTETGG...	..TKEYDID MAFFDNRSAA
C68	MKPCYGSFAK	PTNKEGGQAN VKTGTGT...	..TKEYDID MAFFDNRSAA
Pan5	MKPCYGSFAK	PTNKEGGQAN VKTETGG...	..TKEYDID MAFFDNRSAA
Pan7	MKPCYGSFAK	PTNKEGGQAN VKTETGG...	..TKEYDID MAFFDNRSAA
hAd41	MLPCYGSYAK	PTNEKGGQAS LITNGTD...	..QTLTSDVN LQFFALPST.
hAd40	MQPCYGSYAR	PTNEKGGQAK LVKNDDN...	..QTTTTNVG LNFFTTATET
	301		350
hAd12	A.....N	TAQVVFYTED VNLEMPDTHL	VFKPTVTNGT IASESLLGQQ
hAd3	A.....GAL	APEIVLYTEN VNLETPDSHV	VYKPETSNN. ..SHANLGQQ
hAd7	D.....AF	SPEIVLYTEN VNLETPDSHV	VYKPGTSDDN ..SHANLGQQ
hAd11	T.....NL	SPKIVMYAEN VNLETPDTHV	VYKPGTEDTS ..SEANLGQQ
hAd21	A.....NF	SPKIVMYAEN VDLETPDTHV	VYKPGTSEES ..SHANLGQQ
hAd34	S.....EL	KPKIVMYAEN VDLESPDTHV	VYKPGVSDAS ..SETNLGQQ
hAd35	T.....NF	SPKIVMYAEN VGLETPDTHV	VYKPGTEDTS ..SEANLGQQ
C1TGL	KPKIVMYAEN VDLETPDTHV	VYKPGASDAS ..SHANLGQQ
hAd1	MN...ERNAV	QPSIVLYSED VNMETPDTHI	SYKPSKTDEN ..SKAMLGQQ
hAd2	NDR...QGNAT	KPKVVLYSED VNLETPDTHL	SYKPGKGDEN ..SKAMLGQQ
hAd5	AG...NGDNL	TPKVVLYSED VDIETPDTHI	SYMPTIKEGN ..SRELMGQQ
ChAd3	RN...EANNI	QPKLVLYSED VHMETPDTHL	SYKPAKSDDN ..SKIMLGQQ
ChAd11	RN...EANNI	QPKLVLYSED VHMETPDTHL	SYKPTKSDDN ..SKVMLGQQ
ChAd17	RN...ETNNI	QPKLVLYSED VHMETPDTHL	SYKPAKSDDN ..SKIMLGQQ
ChAd19	NQR...EGNDT	KPKVVLYSED VHMETPDTHI	SYKPTKSDDN ..SKVMLGQQ
ChAd20	AG...NSDNP	TPKVVLYSED VNLETPDTHI	SYMPTNNETN ..SRELLGQQ
hAd48	TGGNGTNVNF	KPDMIMYAEN VNLETPDTHI	VYKPGKEDAS ..SESNLTOQ
ChAd4	AGNG.TNVND	DPDMVMYTEN VNLETPDTHI	VYKPGTSDDS ..SKVNLCOQ
ChAd5	AGNG.TNVND	DPDMVMYTEN VNLETPDTHI	VYKPGTSDDS ..SKVNLCOQ

Fig. 31D


```

ChAd7  AGNG..TNVND DPDMVMYTEN VNLETPDTHI VYKPGTSDDS ..SEVNLCQQ
ChAd16 VTG...GTEEY KADIVMYTEN TYLETPDTHV VYKPGKDNDS ..SKINLVQQ
Pan6   VN...GQDEY KADIVMYTEN TYLETPDTHV VYKPGKDAS ..SEINLVQQ
hAd4   A.....NY DPDIVMYTEN VDLQTPDTHI VYKPGKEDTS ..SESNLGQQ
hAd16  A.....NY DPDIVMYTEN VDLQTPDTHI VYKPGTEDTS ..SESNLGQQ
ChAd6  A.....SF SPELVLYTEN VDLDTPDTHI IYKPGTDETS ..SSFNLGQQ
ChAd9  T.....AGL TPEIVLYTEN VDLETPDTHI VYKAGTDDSS ..SSINLGQQ
ChAd10 A.....AGL APEIVLYTEN VDLETPDTHI VYKAGTDDSS ..SSINLGQQ
C68    A.....AGL APEIVLYTEN VDLETPDTHI VYKAGTDDSS ..SSINLGQQ
Pan5   AA.....GL APEIVLYTEN VDLETPDTHI VYKAGTDDSS ..SSINLGQQ
Pan7   AA.....GL APEIVLYTEN VDLETPDTHI VYKAGTDDSS ..SSINLGQQ
hAd41  .....PN EPKAVLYAEN VSIEAPDTHL VYKPDVAQGT ISSADLLTQQ
hAd40  A.....NF SPKVVLVYSED VNLEAPDTHL VFKPDVNGTS ..AELLGQQ

```

351

400

```

hAd12  AAPNRANYIA FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd3   AMPNRPNYIG FRDNFVGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd7   AMPNRPNYIG FRDNFVGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd11  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd21  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd34  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd35  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
C1     SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd1   AMPNRPNYIA FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd2   SMPNRPNYIA FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd5   SMPNRPNYIA FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd3  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd11 AMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd17 SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd19 SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd20 AMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd48  SMPNRPNYIG FRDNFVGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd4  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd5  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd7  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd16 SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
Pan6   SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd4   AMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd16  AMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd6  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd9  SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
ChAd10 SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
C68    AMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
Pan5   SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
Pan7   SMPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd41  AAPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL
hAd40  AAPNRPNYIG FRDNFIGLMY YNSTGNMGVL AGQASQLNAV VDLQDRNTEL

```

401

450

```

hAd12  SYQLMLDALG DRTRYFSLWN SAVDSYDPDV RVIENHGVED ELPNYCFPLS
hAd3   SYQLLLDSL G DRTRYFSMWN QAVDSYDPDV RRIENHGIED ELPNYCFPLN
hAd7   SYQLLLDSL G DRTRYFSMWN QAVDSYDPDV RRIENHGIED ELPNYCFPLD
hAd11  SYQLLLDSL G DRTRYFSMWN QAVDSYDPDV RVIENHGVED ELPNYCFPLD
hAd21  SYQLLLDSL G DRTRYFSMWN QAVDSYDPDV RRIENHGVED ELPNYCFPLD
hAd34  SYQLLLDSL G DRTRYFSMWN QAVDSYDPDV RVIENHGVED ELPNYCFPLD
hAd35  SYQLLLDSL G DRTRYFSMWN QAVDSYDPDV RVIENHGVED ELPNYCFPLN

```

Fig. 31E

C1	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RVIENTHGVED	ELPNYCFPLD
had1	SYQLLLDSIG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
had2	SYQLLLDSIG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
had5	SYQLLLDSIG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
ChAd3	SYQLLLDSMG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
ChAd11	SYQLLLDSMG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
ChAd17	SYQLLLDSMG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
ChAd19	SYQLLLDSMG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
ChAd20	SYQLLLDSMG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGTED	ELPNYCFPLG
had48	SYQLLLDSLG	DRTRYFSMWN	SAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
ChAd4	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
ChAd5	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
ChAd7	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
ChAd16	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
Pan6	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
had4	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLN
had16	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLN
ChAd6	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLN
ChAd9	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLN
ChAd10	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLN
C68	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
Pan5	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
Pan7	SYQLLLDSLG	DRTRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLD
had41	SYQLMLDALG	DRSRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLG
had40	SYQLMLDALG	DRSRYFSMWN	QAVDSYDPDV	RIIENHGVED	ELPNYCFPLN

451					500
had12	AVG.EIKNYK	GIKPDNG...	..GGGGWTAD	N.TVSEANH	GIGNIAAMEI
had3	GIG.PGHTYQ	GIKVKTD...	..DTNGWEKD	A.NVAPANEI	TIGNNLAMEI
had7	GIG.PAKTYQ	GIKSK.....	..DNGWEKD	D.NVSKSNEI	AIGNNQAMEI
had11	GIG.VPTTSYK	SIVPNGD...	..NAPNWKEP	..EVNGTSEI	GQGNLFAMEI
had21	GVGVPISYK	IIIEPNG....	..QGADWKEP	..DINGTSEI	GQGNLFAMEI
had34	GVGPPQDSYK	EIKPNG....	..DQSTWTNV	..DPNGSSQL	AKGNPFAMEI
had35	GIGVPTTSYK	SIVPNGE...	..DNNNWKEP	..EVNGTSEI	GQGNLSAMEI
C1	GVGPRDSYK	GIETNGD...	..ENTTWKD	L.DPNGISEL	AKGNPFAMEI
had1	GIG.VTDYQ	GIKSNGNG..	..NPQNWTKN	D.DFAARNEI	GVGNPFAMEI
had2	GIG.VTDYQ	AIKANGNGAG	DNGNTTWTKD	E.TFATRNEI	GVGNPFAMEI
had5	GVI.NTETLT	KVKPKTG...	..QENGWEKD	ATEFSKNEI	RVGNPFAMEI
ChAd3	GIG.VTDYQ	AVKTNNNGNG	..GQVTWTKD	E.TFADRNEI	GVGNPFAMEI
ChAd11	GIG.VTDYQ	AVKTNNNGNG	..GQVTWTKD	E.TFAERNEI	GVGNPFAMEI
ChAd17	GIG.VTDYQ	AVKTNNNGNG	..GQVTWTKD	E.TFADRNEI	GVGNPFAMEI
ChAd19	GIG.VTDYQ	VIKT.NGNGQ	..ADPTWEKD	T.EFADRNEI	GVGNPFAMEI
ChAd20	GVI.NTETFT	KVKPKAAQ..	..DAQWEKD	S.EFSKNEI	RVGNPFAMEI
had48	GAG.TNAVYQ	GVKVKTT...	..NNTWEKD	T.AVSEHNQI	CKGNVYAMEI
ChAd4	GAG.TNSVYQ	GVKPKTDN..	..GNDQWETD	S.TVSSHQI	CKGNIYAMEI
ChAd5	GAG.TNSVYQ	GVKPKTDN..	..GNDQWETD	S.TVSSHQI	CKGNIYAMEI
ChAd7	GAG.TNSVYQ	GVKPKTDN..	..GNDQWETD	S.TVSSHQI	CKGNIYAMEI
ChAd16	GSG.TNAAYQ	GVKVKNGQDG	..DVESEWEKD	D.TVAARNQL	CKGNIFAMEI
Pan6	GSG.TNAAYQ	GVKVKNGQDG	..DVESEWEND	D.TVAARNQL	CKGNIFAMEI
had4	GVG.LTDTYQ	GVKVKTDVA..	..GSEKWDKD	DTTVSTANEI	HVGPNPFAMEI
had16	GVG.FTDTYQ	GVKVKTDVA..	..GSEKWDKD	DTTVSTANEI	HVGPNPFAMEI
ChAd6	GVG.FTDTYQ	GIKVKTTNNG	TANATEWESD	T.SVNNANEI	AKGNPFAMEI
ChAd9	AVG.RTNSYQ	GIKPNGG...	..DPATWAKD	E.SVNDANEL	GKGNPFAMEI
ChAd10	AVG.RTDTYQ	GIKANGA...	..DQTTWTKD	D.TVNDANEL	GKGNPFAMEI
C68	AVG.RTDTYQ	GIKANGT...	..DQTTWTKD	D.SVNDANEI	GKGNPFAMEI
Pan5	AVG.RTDTYQ	GIKANGA...	..DQTTWTKD	D.TVNDANEL	GKGNPFAMEI

Fig. 31F

Pan7 AVG.RTDTYQ GIKANGD... ..NQTTWTKD D.TVNDANEL GKGNPFAMEI
 hAd41 GSA.ATDTYS GIKAN.....GQTWTAD DNYADRGAEI ESGNIFAMEI
 hAd40 GQG.ISNSYQ GVKTDN....G.TNWSQN NTDVSSNNEI SIGNVFAMEI

501

550

hAd12 NLQANLWRSF LYSNVGLYLP DDLKYTPGNI KLPDNKNTYE YMNGRVTAPG
 hAd3 NIQANLWRSF LYSNVALYLP DVYKYTPPNI TLPTNTNTYE YMNGRVVSPS
 hAd7 NIQANLWRSF LYSNVALYLP DVYKYTPPNI TLPANTNTYE YMNGRVVSPS
 hAd11 NLQANLWRSF LYSNVALYLP DSYKYTPSNV TLPENKNTYD YMNGRVVPPS
 hAd21 NLQANLWRSF LYSNVALYLP DSYKYTPANV TLPTNNNTYD YMNGRVVPPS
 hAd34 NLQANLWRSF LYSNVALYLP DSYKYTPSNV TLPENKNTYD YMNGRVVPPS
 hAd35 NLQANLWRSF LYSNVALYLP DSYKYTPSNV TLPENKNTYD YMNGRVVPPS
 C1 NIQANLWRSF LYSNVALYLP DSYKYTPPNI TLPENKNTYD YMNGRVVPPS
 hAd1 NLNANLWRNF LYSNIALYLP DKLKYTPPNI EISPNPNSYD YMNKRUVAPG
 hAd2 NLNANLWRNF LYSNIALYLP DKLKYNPTNV EISDNPNNTYD YMNKRUVAPG
 hAd5 NLNANLWRNF LYSNIALYLP DKLKYSPSNV KISDNPNNTYD YMNKRUVAPG
 ChAd3 NLSANLWRNF LYSNVALYLP DKLKYNPSNV DISDNPNNTYD YMNKRUVAPG
 ChAd11 NLNANLWRNF LYSNVALYLP DKLKYNPSNV DISDNPNNTYD YMNKRUVAPG
 ChAd17 NLSANLWRNF LYSNVALYLP DKLKYNPSNV DISDNPNNTYD YMNKRUVAPG
 ChAd19 NLNANLWRNF LYSNVALYLP DKLKYNPSNV DISDNPNNTYD YMNKRUVAPG
 ChAd20 NLNANLWRNF LYSNVALYLP DKLKYTPSNV QISNNPNSYD YMNKRUVAPG
 hAd48 NLQANLWRSF LYSNVALYLP DSYKYTPANV TLPTNTNTYE YMNGRVVAPS
 ChAd4 NLQANLWRSF LYSNVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVPPS
 ChAd5 NLQANLWRSF LYSNVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVPPS
 ChAd7 NLQANLWRSF LYSNVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVPPS
 ChAd16 NLQANLWRSF LYSNVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVPPS
 Pan6 NLQANLWRSF LYSNVALYLP DSYKYTPPNI TLPTNTNTYD YMNGRVTPPS
 hAd4 NIQANLWRNF LYANVALYLP DKYKYTPANI TLPTNTNTYE YMNGRVVAPS
 hAd16 NIQANLWRSF LYSNVALYLP DSYKYTPSNV TLPENKNTYD YMNGRVVPPS
 ChAd6 NIQANLWRNF LYANVALYLP DSYKYTPANI TLPANTNTYD YMNGRVVAPS
 ChAd9 NIQANLWRNF LYANVALYLP DSYKYTPANI TLPANTNTYD YMNGRVVAPS
 ChAd10 NIQANLWRNF LYANVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVAPS
 C68 NIQANLWRNF LYANVALYLP DSYKYTPANV TLPTNTNTYD YMNGRVVAPS
 Pan5 NIQANLWRNF LYANVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVAPS
 Pan7 NIQANLWRNF LYANVALYLP DSYKYTPANI TLPTNTNTYD YMNGRVVAPS
 hAd41 NLAANLWRSF LYSNVALYLP DSYKITPDNI TLPENKNTYA YMNGRVAVPS
 hAd40 NLAANLWRSF LYSNVALYLP DSYKITPDNI TLPDNKNTYA YMNGRVAVPS

551

600

hAd12 LVDTYVNIGA RWSPDVMNDV NPFNHHRNAG LRYRSMLLGN GRVVPFHIQV
 hAd3 LVDSYINIGA RWSLDPMNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd7 LVDSYINIGA RWSLDPMNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd11 LVDTYVNIGA RWSLDAMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd21 LVDTYVNIGA RWSLDAMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd34 LVDTYVNIGA RWSLDAMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd35 LVDTYVNIGA RWSLDAMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 C1 LVDTYVNIGA RWSLDAMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd1 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd2 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd5 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 ChAd3 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 ChAd11 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 ChAd17 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 ChAd19 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 ChAd20 LVDCYINLGA RWSLDYMDNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV
 hAd48 LVDAYINIGA RWSLDPMNDV NPFNHHRNAG LRYRSMLLGN GRYVPFHIQV

Fig. 31G

```

ChAd4  LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
ChAd5  LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
ChAd7  LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
ChAd16 LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
Pan6   LVDAYLNIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
hAd4   LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
hAd16  LVDTYVNIGA  RWSLDAMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
ChAd6  LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
ChAd9  LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
ChAd10 LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
C68    LVDSYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
Pan5   LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
Pan7   LVDAYINIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
hAd41  ALDTYVNIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV
hAd40  ALDTYVNIGA  RWSLDPMDNV  NPFNHHRNAG  LRYRSMLLGN  GRYVPFHIQV

```

```

601
hAd12  PQKFFAIRNL  LLLPGSYTYE  WNFRKDVNMI  LQSTLGNDLR  VDGASVRFDN
hAd3   PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  TDGATISFTS
hAd7   PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  TDGATISFTS
hAd11  PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASISFTS
hAd21  PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASISFTS
hAd34  PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASISFTS
hAd35  PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASISFTS
C1     PQKFFAVKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASISFTS
hAd1   PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFDS
hAd2   PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFDS
hAd5   PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFDS
ChAd3  PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFES
ChAd11 PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFES
ChAd17 PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFES
ChAd19 PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFES
ChAd20 PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMV  LQSSLGNDLR  VDGASIKFES
hAd48  PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  VDGASVRFDS
ChAd4  PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASISFTS
ChAd5  PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASISFTS
ChAd7  PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASISFTS
ChAd16 PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASISFTS
Pan6   PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASIAFTS
hAd4   PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASITFTS
hAd16  PQKFFAVKNL  LLLPGSYTYV  WNFRKDVNMV  LQSSLGNDLR  VDGATISFTS
ChAd6  PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASIAFTS
ChAd9  PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASIAFTS
ChAd10 PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASIAFTS
C68    PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASISFTS
Pan5   PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASIAFTS
Pan7   PQKFFAIKSL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  TDGASIAFTS
hAd41  PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  VDGASVRFDS
hAd40  PQKFFAIKNL  LLLPGSYTYE  WNFRKDVNMI  LQSSLGNDLR  VDGASVRFDS

```

```

651
hAd12  IALYANFFPM  AHNTASTLEA  MLRNDTNDQS  FNDYLCAANM  LYPIPANATS
hAd3   INLYATFFPM  AHNTASTLEA  MLRNDTNDQS  FNDYLSAANM  LYPIPANATN
hAd7   INLYATFFPM  AHNTASTLEA  MLRNDTNDQS  FNDYLSAANM  LYPIPANATN
hAd11  INLYATFFPM  AHNTASTLEA  MLRNDTNDQS  FNDYLSAANM  LYPIPANATN
hAd21  INLYATFFPM  AHNTASTLEA  MLRNDTNDQS  FNDYLSAANM  LYPIPANATN

```

Fig. 31H

hAd34	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd35	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
C1	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd1	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd2	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd5	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd3	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd11	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd17	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd19	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd20	ICLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd48	VNLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPAKATN
ChAd4	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd5	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd7	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd16	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
Pan6	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd4	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd16	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd6	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd9	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
ChAd10	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
C68	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
Pan5	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
Pan7	INLYATFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLSAANM	LYPIPANATN
hAd41	INLYANFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLCAANM	LYPIPSNATS
hAd40	INLYANFFPM	AHNTASTLEA	MLRNDTNDQS	FNDYLCAANM	LYPIPANATS

701

750

hAd12	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GTIPYLDGTF
hAd3	IPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd7	IPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd11	IPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd21	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd34	IPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd35	IPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
C1	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd1	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGYDPYYTYS	GSIPYLDGTF
hAd2	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGYDPYYTYS	GSIPYLDGTF
hAd5	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGYDPYYTYS	GSIPYLDGTF
ChAd3	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGFDPYYTYS	GSIPYLDGTF
ChAd11	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGFDPYYTYS	GSIPYLDGTF
ChAd17	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGFDPYYTYS	GSIPYLDGTF
ChAd19	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGFDPYYTYS	GSIPYLDGTF
ChAd20	VPISIPSRNW	AAFRGWAFTR	LTKKETPSLG	SGFDPYYTYS	GSIPYLDGTF
hAd48	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd4	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd5	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd7	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd16	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
Pan6	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd4	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd16	IPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd6	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd9	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
ChAd10	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF

Fig. 31I

C68	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFVYS	GSIPYLDGTF
Pan5	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
Pan7	VPISIPSRNW	AAFRGWSFTR	LKTRETPSLG	SGFDPYFVYS	GSIPYLDGTF
hAd41	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFTYS	GSVPYLDGTF
hAd40	VPISIPSRNW	AAFRGWSFTR	LTKKETPSLG	SGFDPYFTYS	GSVPYLDGTF

	751		800
hAd12	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
hAd3	YLNHTFKKVA	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd7	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd11	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd21	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd34	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVQCENMT
hAd35	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
C1	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd1	YLNHTFKKVA	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
hAd2	YLNHTFKKVA	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
hAd5	YLNHTFKKVA	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
ChAd3	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
ChAd11	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
ChAd17	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
ChAd19	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
ChAd20	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
hAd48	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRSVDGE GYNVAQCENMT
ChAd4	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
ChAd5	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
ChAd7	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
ChAd16	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
Pan6	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd4	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd16	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
ChAd6	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
ChAd9	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
ChAd10	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
C68	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
Pan5	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
Pan7	YLNHTFKKVS	ITFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd41	YLNHTFKKVS	IMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT
hAd40	YLNHTFKKVS	VMFDSSVSWP	GNDRLLPNE FEIKRTVDGE GYNVAQCENMT

	801		850
hAd12	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVVDTEYK
hAd3	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVNYT
hAd7	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVNYT
hAd11	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVNYK
hAd21	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVINYK
hAd34	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVINYK
hAd35	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVINYK
C1	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDEVINYK
hAd1	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK
hAd2	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK
hAd5	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK
ChAd3	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK
ChAd11	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK
ChAd17	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK
ChAd19	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY SFFRNFQPM RQVDDTKYK

Fig. 31J

ChAd20	KDWFLVQMLA	NYNIGYQGFY	IPESYKDRMY	SFFRNFPQMS	RQVVDQTKYK
hAd48	KDWFLVQMLS	HYNIGYQGFH	VPEGYKDRMY	SFFRNFPQMS	RQVVDEINYNK
ChAd4	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
ChAd5	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
ChAd7	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
ChAd16	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
Pan6	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
hAd4	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
hAd16	KDWFLVQMLA	NYNIGYQGFY	IPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
ChAd6	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
ChAd9	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
ChAd10	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
C68	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
Pan5	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
Pan7	KDWFLVQMLA	HYNIGYQGFY	VPEGYKDRMY	SFFRNFPQMS	RQVVDEVNYSK
hAd41	KDWFLVQMLS	HYNIGYQGFY	VPESYKDRMY	SFFRNFPQMS	RQVVNTTTYK
hAd40	KDWFLVQMLS	HYNIGYQGFH	VPESYKDRMY	SFFRNFPQMS	RQVVDTTYK

851					900
hAd12	NYKKVTVEFQ	HNNSGFVGYL	GPTMREGQAY	PANYPYPLIG	QTAVESITQK
hAd3	DYKAVTLPYQ	HNNSGFVGYL	APTMRQGEPY	PANYPYPLIG	TTAVKSVTQK
hAd7	DYKAVTLPYQ	HNNSGFVGYL	APTMRQGEPY	PANYPYPLIG	TTAVKSVTQK
hAd11	DFKAVAIPYQ	HNNSGFVGYM	APTMRQGQPY	PANYPYPLIG	TTAVNSVTQK
hAd21	DYKAVAVPYQ	HNNSGFVGYM	APTMRQGQAY	PANYPYPLIG	TTAVTSVTQK
hAd34	DFKAVAIPYQ	HNNSGFVGYM	APTMRQGQPY	PANYPYPLIG	TTAVNSVTQK
hAd35	DFKAVAIPYQ	HNNSGFVGYM	APTMRQGQPY	PANYPYPLIG	TTAVNSVTQK
C1	DYKAVAVPYQ	HNNSGFVGYM	APTMRQGQAY	PANYPYPLIG	TTAVTSVTQK
hAd1	DYQQVGILHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
hAd2	DYQQVGILHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
hAd5	DYQQVGILHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
ChAd3	DYQEVGIIHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
ChAd11	DYQEVGIIHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
ChAd17	DYQEVGIIHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
ChAd19	DYQEVGIIHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
ChAd20	DYQEVGIIHQ	HNNSGFVGYL	APTMRREGQAY	PANFPYPLIG	KTAVDSITQK
hAd48	DYKAVTLPFQ	HNNSGFTGYL	APTMRQGQPY	PANFPYPLIG	QTAVPSVTQK
ChAd4	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVTSVTQK
ChAd5	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
ChAd7	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVTSVTQK
ChAd16	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
Pan6	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
hAd4	DYQAVTLPYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVTSVTQK
hAd16	DYKAVTLPYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	TTAVKSVTQK
ChAd6	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
ChAd9	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
ChAd10	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
C68	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVTSVTQK
Pan5	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
Pan7	DYQAVTLAYQ	HNNSGFVGYL	APTMRQGQPY	PANYPYPLIG	KSAVASVTQK
hAd41	EYQNVTLPFQ	HNNSGFVGYM	GPTMREGQAY	PANYPYPLIG	QTAVPSLTQK
hAd40	EYQNVTLPFQ	HNNSGFVGYM	GPAIREGQAY	PANYPYPLIG	QTAVPSLTQK

901					950
hAd12	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMD
hAd3	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMD
hAd7	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMD

Fig. 31K

hAd11	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMDE
hAd21	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
hAd34	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMDE
hAd35	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMDE
C1	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
hAd1	KFLCDRTLWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
hAd2	KFLCDRTLWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
hAd5	KFLCDRTLWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
ChAd3	KFLCDRTLWR	IPFSSNFMSM	GALSDLGQNL	LYANSAHALD	MTFEVDPMDE
ChAd11	KFLCDRTLWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
ChAd17	KFLCDRTLWR	IPFSSNFMSM	GALSDLGQNL	LYANSAHALD	MTFEVDPMDE
ChAd19	KFLCDRTLWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
ChAd20	KFLCDRTLWR	IPFSSNFMSM	GALSDLGQNL	LYANSAHALD	MTFEVDPMDE
hAd48	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMDE
ChAd4	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
ChAd5	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
ChAd7	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
ChAd16	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
Pan6	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
hAd4	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
hAd16	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNL	LYANSAHALD	MTFEVDPMDE
ChAd6	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
ChAd9	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
ChAd10	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
C68	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
Pan5	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
Pan7	KFLCDRVMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MNFEVDPMDE
hAd41	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMDE
hAd40	KFLCDRTMWR	IPFSSNFMSM	GALTDLGQNM	LYANSAHALD	MTFEVDPMDE

951

990

hAd12	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd3	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd7	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd11	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd21	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd34	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd35	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
C1	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd1	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd2	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd5	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd3	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd11	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd17	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd19	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd20	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd48	PTLLYVLFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd4	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd5	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd7	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd16	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
Pan6	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd4	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
hAd16	PTLLSLVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT
ChAd6	STLLYVVFV	FDVVRHQPH	RGVIEAVYLR	TPFSAGNATT

Fig. 31L

ChAd9	STLLYVVFEV	FDVVRVHQP	H	RGVIEAVYLR	TPFSAGNATT
ChAd10	STLLYVVFEV	FDVVRVHQP	H	RGVIEAVYLR	TPFSAGNATT
C68	STLLYVVFEV	FDVVRVHQP	H	RGVIEAVYLR	TPFSAGNATT
Pan5	STLLYVVFEV	FDVVRVHQP	H	RGVIEAVYLR	TPFSAGNATT
Pan7	STLLYVVFEV	FDVVRVHQP	H	RGVIEAVYLR	TPFSAGNATT
hAd41	PTLLYVLF	FDVVRIHQ	H	RGVIEAVYLR	TPFSAGNATT
hAd40	PTLLYVLF	FDVVRIHQ	H	RGVIEAVYLR	TPFSAGNATT

ITR0048PV

148/153

SEQ ID NO: 26

AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTG

SEQ ID NO: 27

ATGGAATTCGTTTAAACCATCATCAATAATATACCTC

SEQ ID NO: 28

CGCTGGCACTCAAGAGTGGCCTC

SEQ ID NO: 29

ATGAAGCTTGTTTAAACCCAT CATCAATAATATACCT

SEQ ID NO: 30

ATCTAGACAGCGTCCATAGCTTACCG

SEQ ID NO: 31

ATGCTACGTAGCGATCGCGTGAGTAGTGTTTGGGGGTGGGTGGG

SEQ ID NO: 32

TAGGCGCGCCGCTTCTCCTCGTTCAGGCTGGCG

SEQ ID NO: 33

GATCTAGTTAGTTTAAACGAATTCGGATCTGCGACGCG

SEQ ID NO: 34

TTCGATCATGTTTAAACGAAATTAAGAATTCGGATCC

SEQ ID NO: 35)

TATTCTGCGATCGCTGAGGTGGGTGAGTGGGCG

SEQ ID NO: 36

TAGGCGCGCCCTTAAACGGCATTGTGTTGGAG

SEQ ID NO: 37

CGTCTAGAAGACCCGAGTCTTACCAGT

SEQ ID NO: 38

CGGGATCCGTTTAAACCATCATCAATAATATACCTTATT

SEQ ID NO: 39

ATGGAATTCGTTTAAACCATCATCAATAATATACCTT

SEQ ID NO: 40

ATGACGCGATCGCTGATATCCTATAATAATAAACGCAGACTTTG

SEQ ID NO: 45

TGTCCTACCARCTCTTGCTTGA

SEQ ID NO: 46

GTGGAARGGCACGTAGCG

Fig. 32

149/153

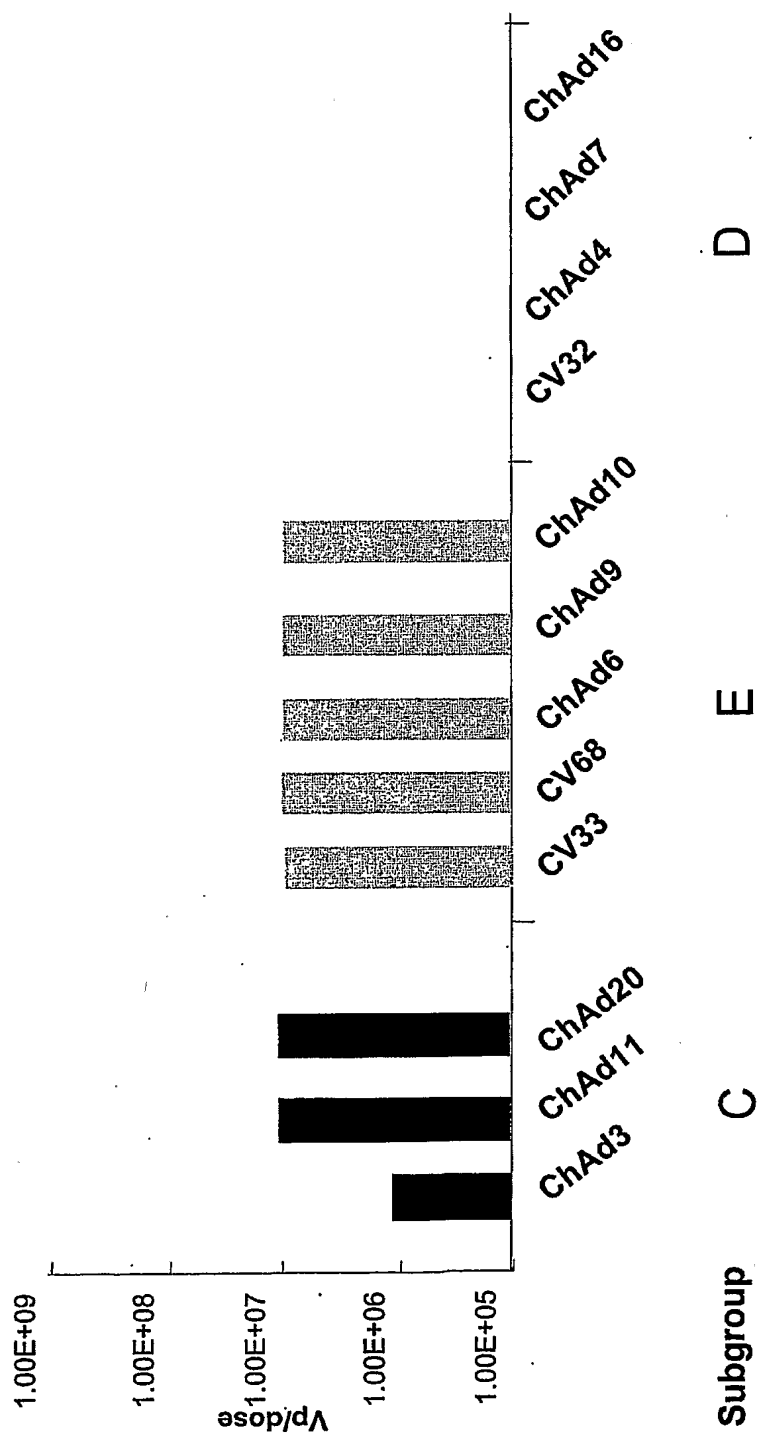


Fig. 33

150/153

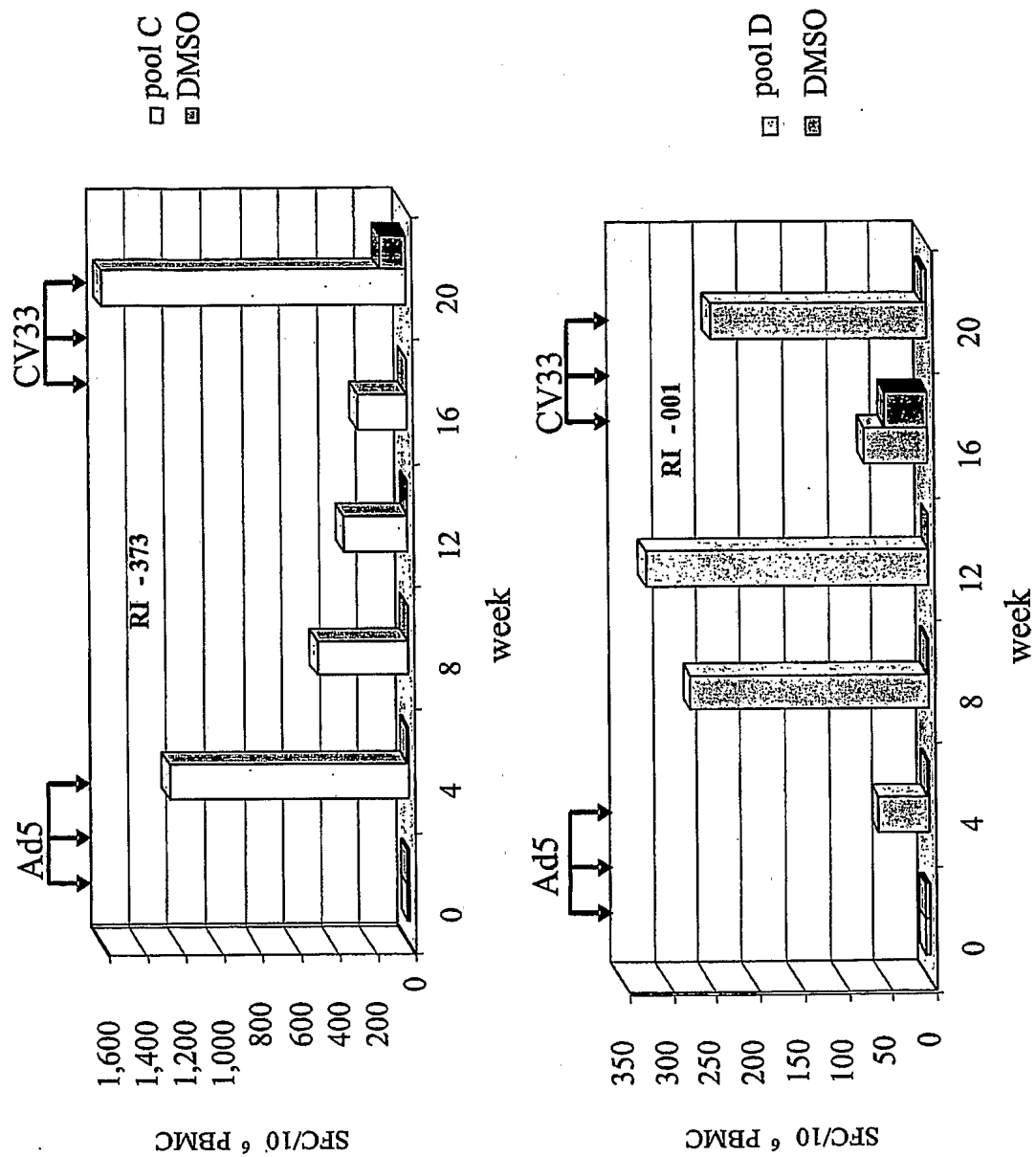


Fig. 34

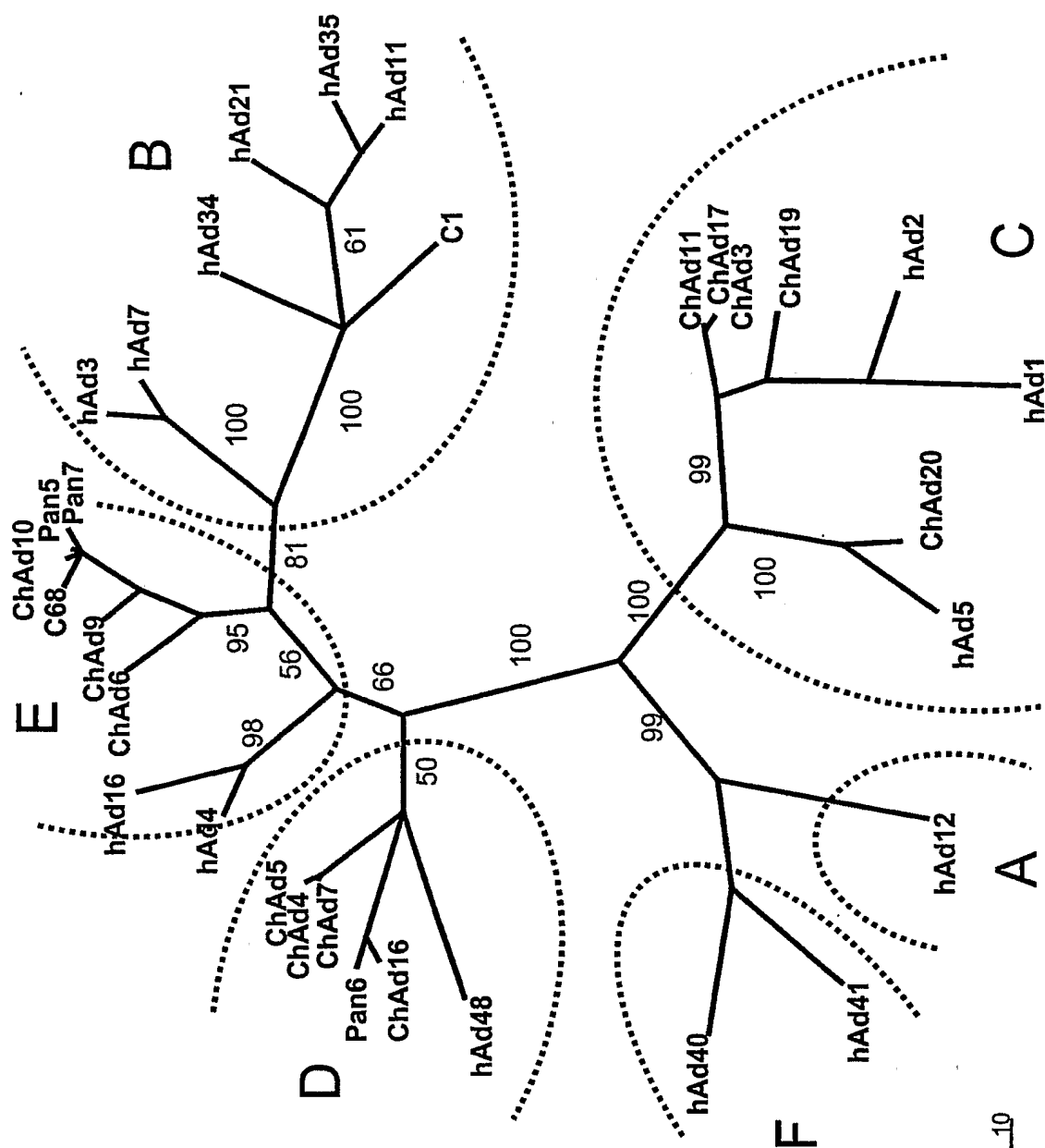


Fig. 35

Pre-immunization with 10^{10} vp Ad5wt at wk 0, 2
 Immunization with MRKAd5 or ChAd3gag at wk 4
 Geomean (n=5) ELISpot responses to gag 9mer

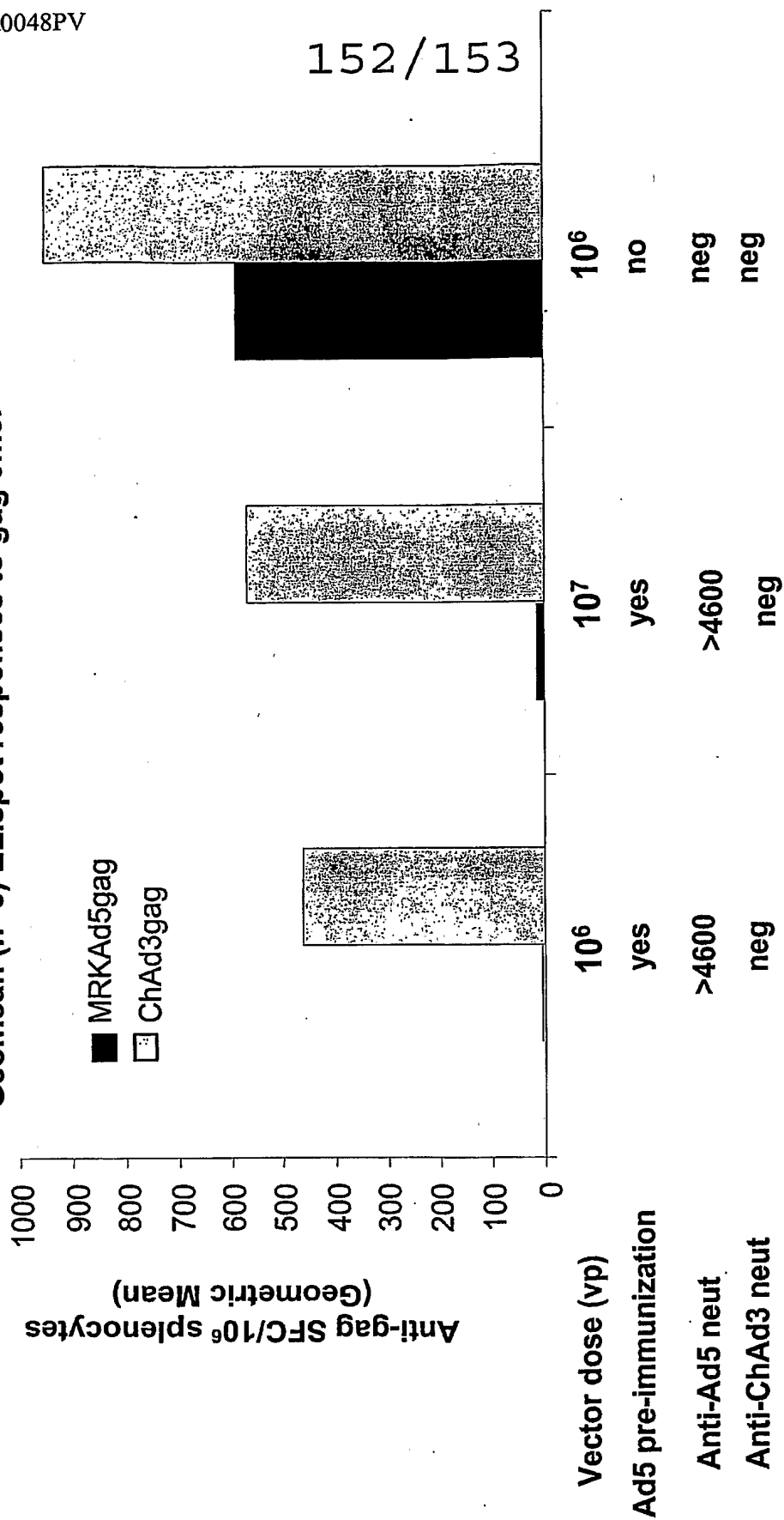


Fig. 36

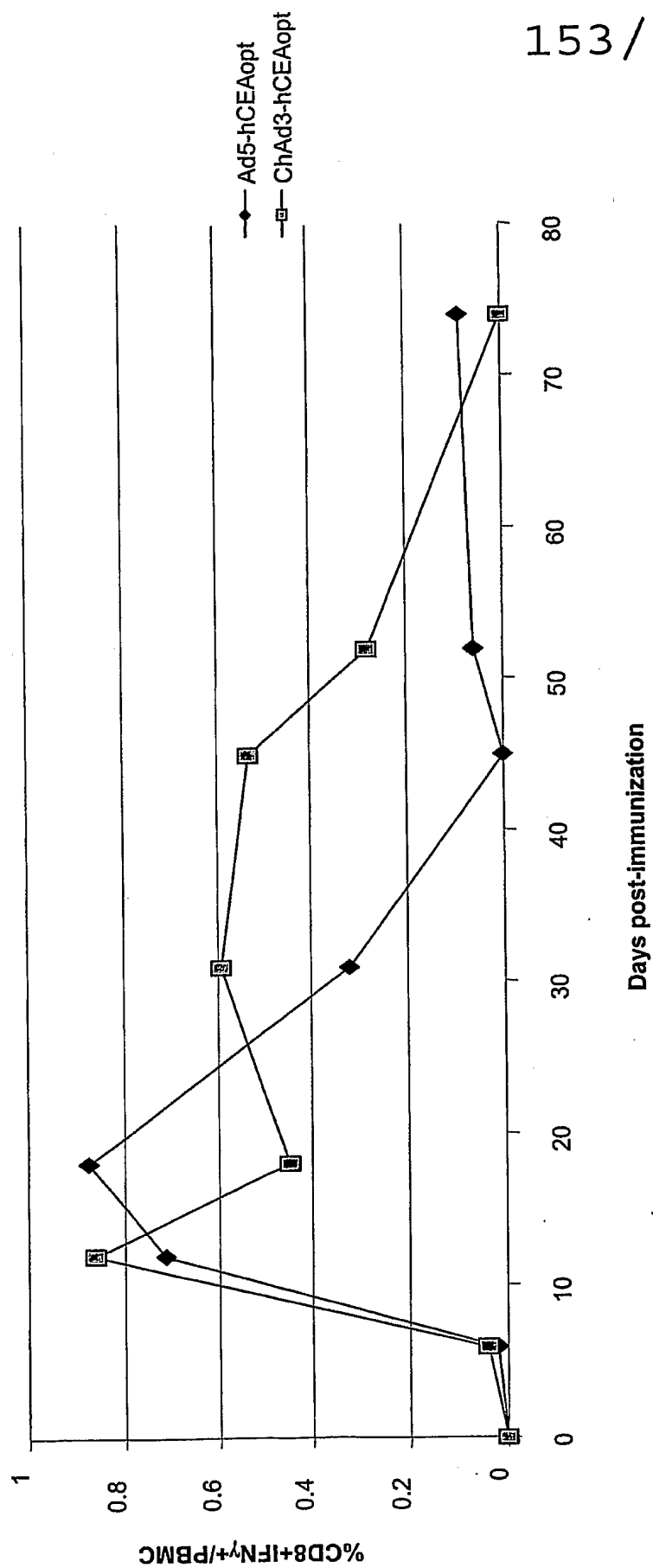


Fig. 37